This manuscript was published as:

Schmalz, V. J., & Cornillie, F. (2022). Towards truly intelligent and personalized ICALL systems using Fluid Construction Grammar. In J. Colpaert, Y. Wang, & G. Stockwell (Eds.), *Proceedings of the XXIst International CALL Research Conference* (pp. 169–179). Castledown Publishers. https://doi.org/10.29140/9781914291050-24

Veronica Juliana Schmalz, Frederik Cornillie

ITEC, imec research group at KU Leuven, Kortrijk, Belgium

veronicajuliana.schmalz@kuleuven.be, frederik.cornillie@kuleuven.be

Towards truly intelligent and personalized ICALL systems using Fluid Construction Grammar

Bio data

Veronica Juliana Schmalz is a PhD student working on the development of a computational model for the usage-based acquisition of modular constructions and grammatical categories. Apart from computational linguistics and AI, she is also interested in language acquisition and development, multilingualism, cognitive linguistics and assessment of learners' competences.

Frederik Cornillie (Ph.D., KU Leuven) is research and valorization manager in educational technology at KU Leuven and at the strategic research institute imec in Belgium. His main research interest in CALL is the intersection of tutorial CALL and task-based language learning, and more specifically how mechanics of games such as feedback can support instructed language learning. He is also passionate about the design and development of CALL applications.

Abstract

Intelligent Computer-Assisted Language Learning (ICALL) aims to design effective systems for the analysis of learners' production in a target language ensuring both successful learning and motivated learners. Most of the existing systems, however, focus extensively on the form rather than on the meaning of language. To create effective systems facilitating personalized language learning both form and meaning should be considered. The reason behind this is that language is a continuous flow of information passing from one user or agent to another, both during comprehension and production. This becomes even more relevant in the case of second or foreign languages (L2), where certain linguistic choices may be dictated by inexact form-meaning links construed by the learner. In this research project, we focus on the analysis of the spoken production of adult learners of German, taking argument and information structure as a use case. We use Fluid Construction Grammar as a formalism, which captures relevant linguistic aspects at both the syntactic level (form) and the semantic level (meaning). Its particularity lies in the possibility of closely monitoring bidirectional form-meaning interactions starting from constructions of different nature modeled in an extensively customizable way. Our work is in progress, and we focus on ways to provide helpful feedback on meaning. German displays a rather articulated grammar and obtaining insights not only on its formal but also on its semantic correctness could offer important steps forward for Intelligent CALL.

The design of computational systems for Intelligent CALL that can effectively support L2 learners in personalized learning requires a grammatical framework that is computationally effective and offers linguistic and acquisitional perspicuity (Schulze & Penner, 2008).

Since human language strictly depends on users and on contexts, we need a system that:

- models the learners' productions in a highly detailed and context-specific way;
- bidirectionally and simultaneously accesses the form and the meaning of the learners' productions highlighting their intersections;
- compares several productions focusing on different language units, also considering their frequency and logical plausibility.

These desiderata are currently found in Fluid Construction Grammar (FCG) (Steels, 2017), a formalism that originated from the intersection of linguistics and AI methods. FCG exploits pattern-finding, generalization and specialization principles (Van Eecke, 2018) to model grammars in the form of constructions, or bidirectional form-meaning pairs. FCG can formally represent elementary linguistic structures, such as lexical units or nominal phrases, and more complex ones, such as idioms. Its application with a free and user-friendly editor (https://www.fcg-net.org/download/) on datasets like ours, containing L2 learners' productions (Baten & Cornillie, 2019), allows us to eventually reach a level of personalization in intelligent tutorial CALL as we have not seen so far. This is due to the possibility of engineering customizable constructions from scratch and concentrating on specific aspects, such as the German argument and information structure, along with the case system (van Trijp, 2011) for our current investigation. Moreover, constructions can have added features that allow their generalization to specific productions, and even the detection of formal or semantic errors. Thanks to FCG and the close observation of form-meaning aspects, we can tackle the following challenges related to either the psychological or the ecological dimension of Smart CALL as defined in the conference theme:

- provide learners with personalized feedback going beyond formal inaccuracies so as to improve their effort/reward ratio and mental acceptance of tutorial CALL activities;
- offer additional information to the teachers or the learners' more skilled peers derived from the application of linguistic constructions, creating opportunities for 'distributed scaffolding' (Tabak & Kyza, 2018) co-provided by the teacher and the CALL system in the ecology around the learner;
- deliver linguistic data to the research community for easily accessible semantically annotated corpora, fundamental to the creation of truly intelligent tutors who can efficiently communicate with learners and provide them with the necessary personalized learning inputs and feedback (Beuls, 2013).

Conference paper

Introduction

Since its establishment, one of the principal goals of Intelligent Computer-Assisted Language Learning (ICALL) was the design of effective systems capable of identifying and correcting the errors of language learners, as well as accurately encoding their competence (Melissa et al., 1993). To measure the effectiveness of such systems, it is crucial to take into account their capacity to enhance the learning experience and improve the motivation of learners. Nevertheless, most existing systems mainly focus on the form rather than on the meaning of language (cf. Schulze & Penner, 2008). In order to take a significant step forward towards the development of more efficient and personalized ICALL systems, the correlation between form and meaning in grammar should receive more attention. The reason behind this is the bidirectionality between the two parts, which turn out to be continuously correlated both in language comprehension and formulation (Goldberg, 2003). This appears especially crucial when learning a new language, since learners need to focus on properly understanding and correctly generating utterances in a different language than their native one.

The acquisition of a new language system, namely of its grammar, represents a demanding challenge since each language has a unique way of representing reality through distinct syntactic and semantic units and principles. Monitoring the interactions between these different components in a clear and understandable way proves to be a complex task as well, since they are thoroughly intertwined and subject to variation. Construction grammars come in response to this daunting task. Their name refers to their fundamental element, that is the construction, a more or less complex unit consisting of a single or multiple words with different functions, characterized by a form and a meaning, as well as other syntactic and semantic features attributed to it (Goldberg, 2006). Construction grammars have been previously regarded as a suitable framework for ICALL systems by Schulze and Penner (2008) since they meet the requirements of *computational effectiveness*, *linguistic perspicuity* and *acquisitional perspicuity* introduced by Matthews (1997).

According to usage-based theories, mainly based on general cognitive capacities such as intention reading and pattern finding (Tomasello, 2003), construction grammars can be used to faithfully model the processes of first-language acquisition. Through the years several formalisms have emerged, among which:

- HPSG - Head-driven Phrase Structure Grammar (cf. Copestake & Flickinger, 2000)
- ECG – Embodied Construction Grammar (cf. Bergen & Chang, 2005)
- SBCG – Sign-Based Construction Grammar (cf. Michaelis, 2009)
- FCG – Fluid Construction Grammar (cf. Steels, 2011)

Recent computational implementations have made it possible to investigate in more detail the adequacy of construction grammars to model foreign language learning strategies. Among these, FCG constitutes the only bidirectional framework to computationally implement construction grammars. It originated in the 1990s from the combination of experts' knowledge in the artificial intelligence domain, as well as the cognitive and formal linguistics domain. Initially, it was adopted to model the process of grammar acquisition in artificial agents (Steels at al., 2012). Today, it can be applied to more general or specific domains, such as representing and monitoring the usage-based acquisition of a language's grammar (Doumen et al., 2021) or investigating the case system in German (van Trijp, 2011). Particularly, it enables the mapping between an utterance and its meaning, as well as the opposite from the meaning, which can be expressed in any form, to the utterance. This can be all managed through a user-friendly interface[1] that displays the different engineered constructions individually and their progressive interactions. The uniqueness and strength of FCG lies in its *fluidity*, namely in the possibility of endless customization and the addition of linguistic features. This provides a robust way to parse and analyze in detail linguistic data like ours, as well as to design target corrections and feedback for learners (see Sections 5 and 6).

Case study: Argument and information structure in German

In this research project, we focus on the analysis of 1,487 transcribed and annotated spoken productions of 36 students of German as a second language (L2) from the academic program in Languages and Literature at the University of Ghent, collected in an oral elicited imitation task (Baten & Cornillie, 2019). The learners heard an audio stimulus containing transitive and ditransitive verbs, along with noun and prepositional phrases, which they had to match to one of two displayed pictures with different semantic roles. Subsequently, they needed to produce an oral response describing the selected picture, which should have corresponded to the initially heard utterance's meaning (see Figure 1).

[1] *https://www.fcg-net.org/*

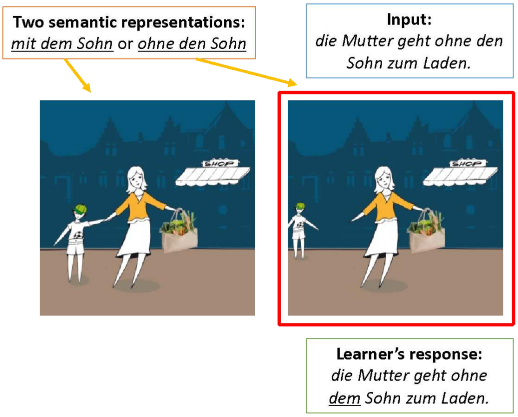

Figure 1. *Two pictures containing different semantic representations of an event are presented to the students. In the left one the mother goes to the shop with the son, while in the right one she does it without the son. The stimulus received by the students concerns the picture on the right. In the example response the student produces a semantically correct utterance, which however contains a formal error in case selection. The preposition 'ohne' requires an accusative and not a dative determiner.*

We are interested in the acquisition of the German argument and information structure. We explore how these can be modeled based on the case system and Abstract Meaning Representation (AMR) rules for semantic roles (cf. Banarescu et al., 2013) in a tailor-made FCG grammar. Exploiting the versatility of FCG and a user-friendly interface, our aim is first to create a grammar capable of a bidirectional representation of the stimuli, namely 48 utterances used to describe the pictures. Once this grammar has been created and tested both in comprehension—namely providing an utterance FCG outputs its meaning (including the argument structure with semantic roles) and in formulation-—namely providing an AMR meaning FCG can generate a desired utterance, we can proceed to model the learners' responses. To do so, we, as grammar engineers, have to implement a series of rules that apply in the case of correct and incorrect responses. In the case of the latter, there are two different approaches to handle students' errors: mal-rules and constraint relaxation. The former consists in the creation of rules based on learners' errors using mal-rules (cf. Sleeman, 1982; Matthews, 1992). The latter relaxes the constraints of the original grammatical rules so that they can also apply in the case of errors (Foth et al., 2005). Since there is no set of rules capable of capturing all the errors expected by learners, the two aforementioned methods can be combined in FCG to detect errors and provide feedback, as well as to allow the processing and understanding of learners' productions (see Sections 5 and 6).

Fluid Construction Grammar (FCG)

Differently from other computational construction grammars, FCG does not have specific implementation principles. Therefore, it is suitable for representing how learners, or in general language users, perceive and utter about the world using grammar. Importantly, FCG relies on the principle of bidirectionality between form and meaning. As a consequence, we can exploit it both to comprehend given utterances and to formulate new ones. Each FCG construction, namely the linguistic unit, needs to have a semantic part related to meaning and a syntactic part related to form.

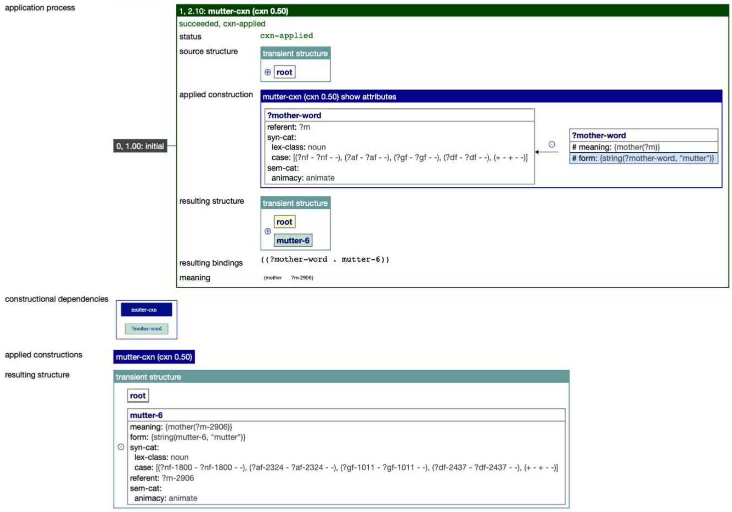

Figure 2. *Lexical construction 'Mutter' with form-meaning information displayed in the application process of constructions and resulting transient structure*

There can be different types of constructions, such as lexical constructions for nouns with a basic meaning and form, along with some basic semantic (e.g. *animacy*) and syntactic (e.g., *lex-class noun*) information (see Figure 2), and phrasal ones that combine lexical constructions for nouns with other units, for example determiners or prepositions (see Figure 3).

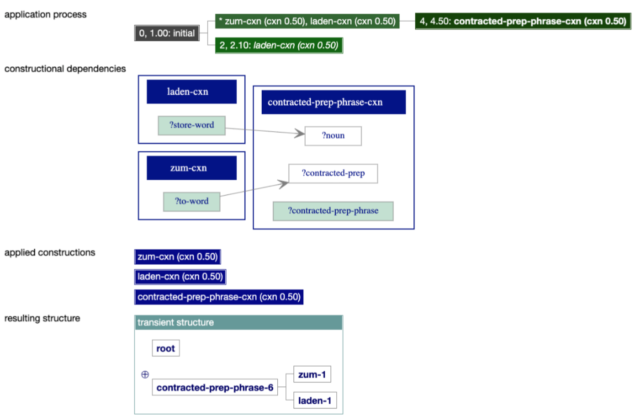

Figure 3. *Comprehension of the prepositional phrase 'zum Laden'. The application process represents how smaller constructions contribute to a larger prepositional phrase construction, based on a dative masculine contracted preposition 'zum' and a masculine noun 'Laden'*

Concomitantly, constructions can map meaning to semantic categories but also syntactic categories to a form, namely a word or phrase, and semantic categories to syntactic units. To represent how constructions are used by a language user or learner, they are arranged in a data structure called *transient structure*. It displays how the processing of a linguistic utterance happens during the application of constructions.

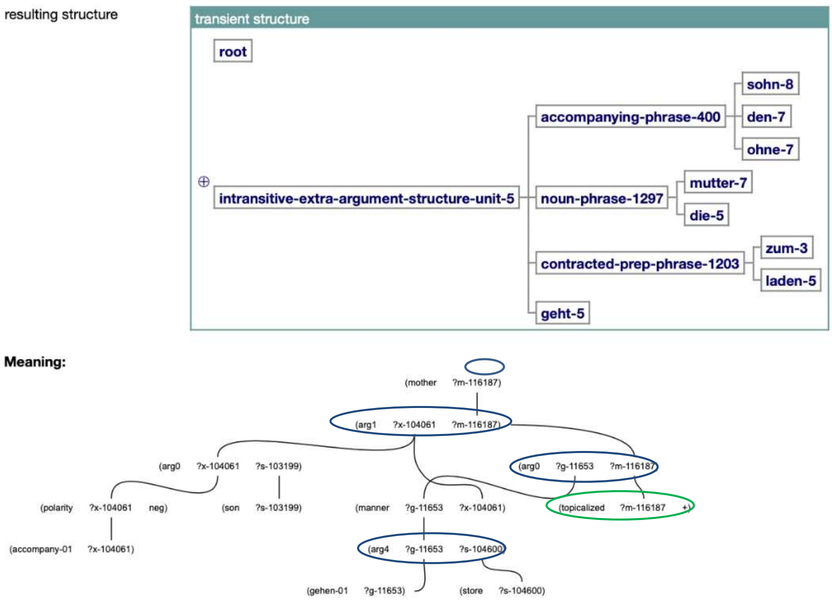

Figure 4. *Resulting transient structure and meaning networks for the comprehension of the utterance 'die Mutter geht ohne den Sohn zum Laden'*

Constructions and transient structures regularly interact with each other via matching and merging processes. For instance, in comprehension an initial transient structure, namely a visual representation of how the form is organized, is built. Afterwards, the matching checks whether the features of existing constructions can apply to the given stimuli. Therefore, if successful, the transient structure receives additional information. The final transient structure is obtained after all constructions have applied. It contains morphological and lexical features as well as data relative to the information and argument structure (see Figure 4).

A computational model for the German grammar

In the last years, FCG has been used to implement different grammars focusing on distinct aspects, for example the Spanish verb conjugation (Beuls, 2013), English quantifiers (Pauw & Hilferty, 2012), Russian verbal aspect (Gerasymova, 2012), Hungarian verbal agreement (Beuls, 2011), Polish negation (Höfer, 2012), L2 Spanish verb morphology (Beuls, 2014), German spatial language (Spranger & Loetzsch, 2011) and German case system (van Trijp, 2011). Since FCG is a dynamic, adaptive and flexible tool, it can be efficiently exploited to represent how languages are continuously influenced and modified by language users' diverse usages and variations. This makes it a perfectly functioning formalism for our project since we are handling learners' productions in a rich and elaborate grammatical system like the German one.

One of the distinctive aspects of German is its rich and articulated case system. It can be a rather complex concept to learn, especially for those learners whose native languages do not exhibit such a range of case options. The information indicating the case is often encapsulated in the determiner or in the suffixes of adjectives, and in some cases also in nouns. However, the presence of syncretism, that is the polyvalence of determiners or suffixes for case, number and gender, can often cause issues in ruling out one single case among the different options and assigning a syntactic role to a given construction. For example, a sentence containing two determined feminine nouns as subject and as object can create confusion, especially when the argument and information structure do not follow the standard order and one is topicalized (e.g., *die Katze* $_{object}$ *sieht* *die Gans* $_{subject}$ *aus der Ferne* – the cat sees the goose from far away).

Moreover, when dealing with prepositions that require a specific case to acquire a certain meaning, the correct application and recognition of the case system is crucial. These aspects of German grammar are present in our students' data. Therefore, our task was to model them using FCG in the most efficient and least ambiguous way possible. To solve the case syncretism issue in computationally modeling German grammar, we decided to follow the feature matrix approach proposed by van Trijp (2011). It is inspired by the distinctive features with binary values used in phonology and based on the application of unification processes. Despite its apparent simplicity, it allows to combine information related to gender, number and case, and compares it across multiple levels and units. In a similar fashion, the matrix is used for the argument structure of verbs. For example, transitive verbs need the nominative case for the agent, or subject, and the accusative for the patient, or object (see Figure 4 - meaning network). The task is made more complex in the case of prepositional phrases with motion verbs and topicalized arguments.

Each verb presents a different argument structure depending on the macro-category to which it belongs according to the formalism of AMR (cf. Banarescu et al., 2013). AMR is a semantic representation that makes use of PropBank framesets (Kingsbury & Palmer 2002) and labeled graphs to analyze and connect different parts of speech in an understandable way both for human users and artificial agents. The goal of AMR is to be an intuitive and easily interpretable, yet consistent and inclusive annotation. Since we want our grammar to be applicable both in formulation and in comprehension, we also added features to rule out the possible alternatives among several prepositional phrases and information structures. In fact, we distinguish contracted and non-contracted prepositional phrases, but also stative, motion, means and accompaniment prepositional phrases which can match specific verbs arguments. On the other hand, for what concerns the information structure, we exploit FCG possibilities to model it separately from the argument structure, highlighting variations from the non-topicalized standard order.

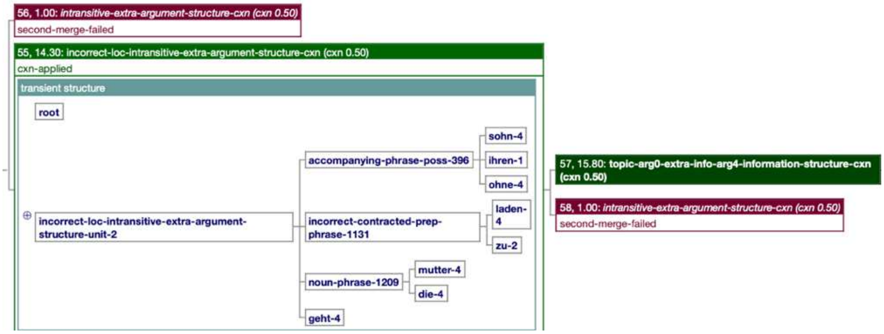

Figure 5. *Constructions' application process in the comprehension of an incorrect intransitive utterance. Red constructions could not be applied, while green ones could. The error here is caused by the undetermined prepositional phrase 'zu Laden' which inhibits the application of the argument structure construction but not of the independent information structure*

Detecting errors

We have engineered an FCG grammar according to a hand-written set of specific constructions related to the stimuli sentences that were provided to the students during the oral elicited imitation task. With this grammar, correct German sentences can be comprehended and automatically generated. We can extend this by adding the variations produced by the learners. These can be of three types:

- formal errors denoting lack of competence in certain grammatical aspects,
- meaning errors signaling incorrect understanding of the stimuli or interpretation of the argument structure,
- variations with respect to the received stimulus but still formally and semantically correct.

When attempting to parse a student response containing one or more of the above variations, the created computational model of German grammar detects differences from the norm. These differences are identified and visually displayed in the FCG web interface (see Figure 5).

For the variations found in our data, contained in 282 sentences, our approach is to implement mal-rules that allow us to go beyond error detection and continue with the comprehension of the students' sentences using our grammar. These are part of a specific construction set that comes into play whenever the pre-existing constructions of the designed grammar fail to apply (see Figure 6). We can also add information containing clarifications about the variations, such as indications on what we would have expected and why these are considered incorrect or improper. This constitutes the basis for providing feedback on meaning, rather than exclusively on form (see Section 6).

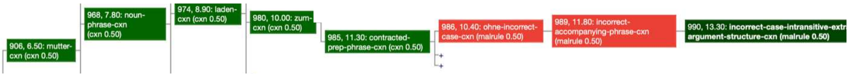

Figure 6. *Example of mal-rules (in orange) in the application process for the comprehension of 'die Mutter geht ohne dem Sohn zum Laden'. In this error case, the dative instead of accusative case is used with the preposition 'ohne'.*

Providing meaningful feedback

Given its bidirectionality and open-endedness, FCG allows us to handle form and meaning errors, as well as variations in detail, apart from simply detecting them. These two possibilities prove its adequacy for the analysis of language learners' production as well as for the design of didactically relevant feedback, tailored to the specifically targeted language aspects and learners' metalinguistic knowledge. For example, when a student utters a sentence such as 'der Doktor verkauft den Clown das Buch', although each noun phrase has a correctly matched determiner, its meaning is incorrect since the AMR required arguments for the ditransitive verb 'verkaufen – to sell' are *arg0* (subject-seller) in nominative, *arg1* (object-sold entity) in accusative and *arg2* (receiver-buyer) in dative. This error is captured by a specific mal-rule that we complemented with additional information concerning the error type and the reason for it. The consequences of the error can also be seen in the resulting transient structure where the arg2 field is empty and there are two possible arg1s (see Figure 7).

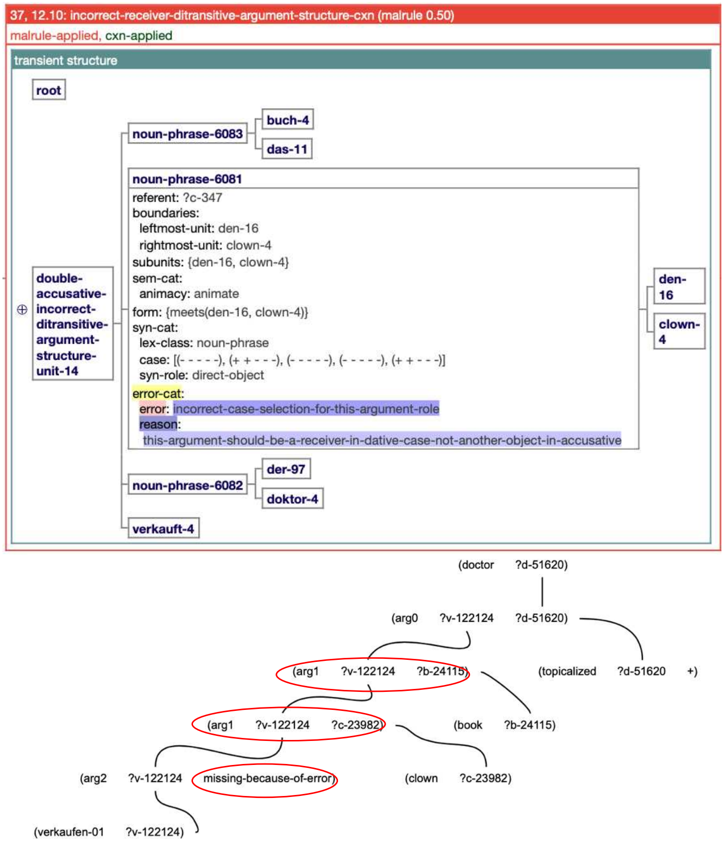

Figure 7. *Incorrect construction with mal-rule application and meaning network for the utterance 'der Doktor verkauft den Clown das Buch' with two accusatives instead of a dative case for arg2.*

In addition to this type of cases, we have cases in which the learner modifies the stimulus sentence more freely. For example, instead of *'die Mutter geht ohne den Sohn zum Laden*' a learner can say *'die Mutter geht ohne ihren Sohn zum Laden*'. Although it deviates from the expected answer, considering the original stimulus, it is not ungrammatical, and the meaning does not differ from the expected response. Therefore, it can be accepted as a variation, and the grammatical constructions can apply effortlessly, as long as the new word used is added to the lexicon.

With these methods we are trying to go beyond error detection and personalize the feedback that learners receive depending on their metalinguistic awareness and competence, as well as on the areas of concern of a given exercise or test in a foreign language.

Conclusion

With this project we are experimenting with the use of a computational construction grammar formalism to represent not only the engineering of the grammar of a correct and established language but also variations on the target language as produced by language learners. Exploiting the possibility of extensive customization, the modularity of constructions and the efficiency of the FCG editor in visualizing the interactions between form and meaning, as well as between argument and information structure, FCG appears to be a promising tool to be used in the study of language acquisition processes, even beyond its original conception for artificial agents. The ability to precisely detect the source and site of errors in the path of application of constructions brings us closer to the reasoning that possibly a learner makes and enables us to provide more in-depth and personalized feedback. In order to do that, we are closely analyzing our dataset, especially the students' responses, searching for generalizable patterns and possibilities to scale up the design and application of constructions. However, this is an uphill climb that requires close collaboration between researchers in computational linguists and language acquisition, teachers, and the learners themselves in order to bring significant improvements and new insights to ICALL.

References

Banarescu, L., Bonial, C., Cai, S., Georgescu, M., Griffitt, K., Hermjakob, U., Knight, K., Koehn, P., Palmer, M., & Schneider, N. (2013). Abstract meaning representation for SemBanking. In Proceedings of the 7th linguistic annotation workshop and interoperability with discourse, 178-186.

Baten, K., & Cornillie, F. (2019). Elicited imitation as a window into developmental stages. *Journal of the European Second Language Association*, *3*(1), 23–34.

Bergen, B. K., & Chang, N. (2005). Embodied construction grammar in simulation- based language understanding. In J-O. Östman & M. Fried (Eds.), *Construction grammar(s): Cognitive and cross-language dimensions* (pp. 147-190). Amsterdam: John Benjamins.

Beuls, K. (2011). Construction sets and unmarked forms: a case study for Hungarian verbal agreement. In L. Steels (Ed.), *Design patterns in fluid construction grammar* (pp. 237–264). Amsterdam: John Benjamins.

Beuls, K. (2013). Towards an agent-based tutoring system for Spanish verb conjugation. [Doctoral dissertation, Vrije Universiteit Brussel].

Beuls, K. (2014) Grammatical error diagnosis in fluid construction grammar: A case study in L2 Spanish verb morphology, *Computer Assisted Language Learning*, *27*(3), 246-260.

Copestake, A., & Flickinger, D. (2000). An open-source grammar development environment and broad-coverage English grammar using HPSG. In Proceedings of the 2nd international conference on language resources and evaluation (LREC 2000), 591–600.

Croft, W. (2001). *Radical construction grammar. Syntactic theory in typological perspective*. Oxford: Oxford University Press.
Doumen, J., Beuls, K. & Van Eecke, P. (2021). A computational model of the co-emergence of item-based constructions and syntactic categories. 11th International Conference on Construction Grammar (ICCG11).
Foth, K., Menzel, W., & Schröder, I. (2005). Robust parsing with weighted constraints. *Natural Language Engineering*, *11*(1), 1–25.
Gerasymova, K. (2012). Expressing grammatical meaning with morphology: A case study for Russian aspect. In L. Steels (Ed.), *Computational issues in fluid construction grammar* (pp.91-112). Berlin: Springer Verlag.
Goldberg, A. E. (2003). Constructions: A new theoretical approach to language. *Trends in Cognitive Sciences*, *7*(5), 219-224.
Goldberg, A. E. (2006). Constructions at work: The nature of generalization in language. Oxford University Press on Demand.
Höfer, S. (2012). Complex declension systems and morphology in Fluid Construction Grammar: A case study of Polish. In L. Steels (Ed.), *Computational issues in Fluid Construction Grammar* (pp.143-177)*.* Berlin: Springer Verlag.
Kingsbury, P.R. & Palmer, M. (2002). From TreeBank to PropBank. LREC.
Matthews, C. (1992). Going AI: foundations of ICALL. *Computer Assisted Language Learning*, *5*(1-2), 13–31.
Matthews, C. (1997). Grammar frameworks in intelligent CALL. *CALICO Journal*, *11*(1). 5–27.
Melissa, H. V., Maisano, R., Alderks, C. & Martin, J. (1993). Parsers in tutors: What are they good for?. *CALICO Journal*, *11*(1), 28–46.
Michaelis, L. A. (2009). Sign-based construction grammar. In B. Heine & H. Narrog (Eds.), *The Oxford handbook of linguistic analysis* (pp. 155-176), Oxford: Oxford University Press.
Pauw, S., & Hilferty, J. (2012). The emergence of quantifiers. E*xperiments in cultural language evolution*, *3*, 277-283.
Schulze, M., & Penner, N. (2008). Construction grammar in ICALL. *Computer Assisted Language Learning*, *21*(5), 427–440.
Sleeman, D. (1982). Inferring (mal) rules from pupils' protocols. *Selected and updated papers from the proceedings of the 1982 European conference on Progress in artificial intelligence*, 30-39.
Spranger, M., & Loetzsch, M. (2011). "Syntactic indeterminacy and semantic ambiguity." In L. Steels (Ed.), *Design patterns in fluid construction grammar* (pp. 237–264). Amsterdam: John Benjamins.
Steels, L. (2011). (Ed.). *Design patterns in Fluid Construction Grammar.* Amsterdam: John Benjamins.
Steels, L., De Beule, J., & Wellens, P. (2012). Fluid construction grammar on real robots. In L. Steels & M. Hild (Eds.), *Language grounding in robots* (pp. 195–213). Berlin, Heidelberg: Springer.
Steels, L. (2017). Basics of fluid construction grammar. *Constructions and Frames*, *9*(2), 178–225.
Tabak, I., & Kyza, E. A. (2018). Research on scaffolding in the learning sciences. International Handbook of the Learning Sciences (pp. 191–200). New York: Routledge. Updated papers from the proceedings of the 1982 European conference in artificial intelligence (pp. 30–39). John Wiley & Sons, Inc.
Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition*. Harvard University Press.
Van Eecke, P. (2018). Generalisation and specialisation operators for computational construction grammar and their application in evolutionary linguistics research. [Doctoral dissertation, Vrije Universiteit Brussel].
van Trijp Remi (2011). Feature matrices and agreement: A case study for German case. In Luc Steels (Ed.), *Design patterns in fluid construction grammar* (pp. 205–235). Amsterdam: John Benjamins.