

# Construction Grammar in ICALL

Mathias Schulze\* and Nikolai Penner

Department of Germanic and Slavic Studies, University of Waterloo, Ontario, Canada

\* e-mail: mschulze@uwaterloo.ca

## *Abstract*

The choice of grammatical framework in ICALL—the branch of CALL that applies artificial intelligence techniques—has important implications for both research and development. Matthews (1993) argued for one “that potentially meshes with SLA” (p. 5) and sketches three criteria which facilitate the crucial decision of selecting a grammatical framework for an ICALL system: computational effectiveness, linguistic perspicuity, acquisitional perspicuity. We will use Matthews’ three adequacy criteria to review recent research in construction grammar and propose its application in ICALL, particularly in projects which involve the building of a student model. Such a student modeling project—Mocha—will be sketched briefly to provide a concrete context for the conceptualization and implementation of construction grammar. This grammatical framework has the potential to help 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.

## *Grammar in ICALL*

Intelligent CALL (ICALL)—the application of artificial intelligence techniques such as natural language processing and student modeling to CALL—used a large variety of grammatical frameworks in the different projects over the last thirty years (Heift & Schulze, 2007, pp. 60-61). The phrase ‘grammatical (or linguistic or syntactic) framework’ is used as a generic term for different formal grammars. The concentration on formal grammars—a description of language which is based on mathematical axioms and principles—is necessary because only such a mathematical approach facilitates an implementation of this grammar in an ICALL system (Schulze, 1999). The focus on grammar has been determined by two factors. First, most of the work done in ICALL thus far concentrated on the processing and the analysis of grammatical structures of learner utterances. And second, the starting point for the discussion in this paper is a paper by Matthews (1993), who already indicates in the title of his paper—“Grammar Frameworks in Intelligent CALL”—that his main focus is grammar. He lists eight grammar frameworks which had been used up until then in ICALL (p. 9): Augmented Phrase Structure frameworks and PATR-II-like environments (Chanier, Pengelly, Twidale, & Self, 1992; Chen & Kurtz, 1989; Labrie & Singh, 1991; Levin, Evans, & Gates, 1991; Sanders, 1991; Schwind, 1990), Augmented Transition Networks (Handke, 1992; Weischedel, Voge, & James, 1978), Lexical Functional Grammar (Feuerman, Marshall, Newman, & Rypa, 1987), Systemic Grammar (Fum, Pani, & Tasso, 1992), Tree Adjoining Grammar (Abeillé, 1992), Incremental Procedural Grammar (Pijls, Daelemans, & Kempen, 1987), Word Grammar (Zähner, 1991), and Preference Semantics (Wilks & Farwell, 1992). Matthews realizes the importance

of the grammatical framework for an ICALL system—not just as a central design decision in the system development, but also a crucial decision for the facilitation of CALL and SLA research. He argues for “choosing a formalism that potentially meshes with SLA” (Matthews, 1993, p. 5), a view which we wholeheartedly support because an appropriate syntactic framework will yield more adequate analyses of learner language and will thus not only improve, for instance, scaffolding and feedback for the learner but also provide new insight into characteristics of learner language and into language learning processes. Matthews sketches three criteria (pp. 9-13) which ought to inform the decision of selecting a grammatical framework for an ICALL system:

- *computational effectiveness* describes the criterion of successful computational implementation of the grammatical framework, which needs to be associated with a formalism (to express the framework in a mathematical and hence computational way) and an efficient parsing algorithm;
- *linguistic perspicuity* refers to the necessary descriptive power of a grammatical framework, a descriptive power which needs to go beyond the adequate description of a few selected linguistic phenomena;
- *acquisitional perspicuity* captures another role of grammar frameworks namely their contribution to the explanation of the acquisition and the development of L2.

With these criteria, he compares rule-based with principle-based frameworks, using Definite Clause Grammar (Pereira & Warren, 1980) as an example of the former and Principles and Parameters Theory (Chomsky, 1986) for the latter. Matthews (1993, 1994) concludes that principle-based frameworks (and consequently principle-based parsing) are the most suitable grammar frameworks for ICALL. In the meantime, these have been used in a number of ICALL projects (Hamel, 1996; L'Haire & Vandeventer Faltin, 2003; Schulze & Hamel, 1998; Vandeventer, 2000, 2001; Vandeventer & Hamel, 2000; Vandeventer Faltin, 2003; Weinberg, Garman, Martin, & Merlo, 1995), but they have certainly not become the dominant framework. In addition to the formal grammars listed by Matthews, after 1993/94 a number of ICALL projects<sup>1</sup> have relied on Generalised Phrase Structure Grammar (GPSG) (Heinecke, Kunze, Menzel, & Schröder, 1998; Menzel & Schröder, 1999) and Head Driven Phrase Structure Grammar (Brocklebank, 1998; Heift, 1998; Heift & Nicholson, 2000; Schulze, 2001). More recently, we have seen the development of cognitive, usage-based grammatical frameworks such as Construction Grammar (CG) in (different branches of) Linguistics. In the following sections, we will discuss the emerging role of this grammatical framework in ICALL.

## *Construction Grammar*

A usage-based view takes grammar to be the cognitive organization of one's experience with language. Aspects of that experience, for instance, the frequency of use of certain constructions or particular instances of constructions, have an impact on representation that are evidenced in speaker knowledge of conventionalized phrases, and in language variation and change. (Bybee, 2006, p. 711)

Construction grammar (CG) is an umbrella term for a number of approaches which all view constructions<sup>2</sup> as the central syntactic unit (for an overview see e.g., Fischer & Stefanowitsch, 2007). CG is thus defined as “a sign-

---

<sup>1</sup> See Heift and Schulze (2007, pp. 52-82) for an overview of ICALL projects.

<sup>2</sup> The term construction is used inside and outside of CG, for instance in structuralism, generative grammar, and corpus linguistics (see Schönefeld, 2006 for a discussion). Here the term is only used to refer to the most essential concept of CG.

based grammatical model that is organized around the notion of GRAMMATICAL CONSTRUCTION as the basic unit of analysis and representation" (Östman & Fried, 2004, p. 12). Unlike most other theories of grammar, it does not form a single unified theory but is rather a "family of loosely connected models" (p. 1). The more popular approaches within this framework are: the Berkeley School (Kay, 1995, 1997, 2002; Kay & Fillmore, 1999);<sup>3</sup> cognitive CG (Goldberg, 1995, 2006); and Radical Construction Grammar (Croft, 2001). The main tenets of CG, accepted by all of its branches, are:

1. the notion of grammatical *construction* as a basic unit of linguistic representation:

A construction is "a form-meaning pair (F, M) where F is a set of conditions on syntactic and phonological form and M is a set of conditions on meaning and use" (Fischer & Stefanowitsch, 2007, p. 5; Lakoff, 1987, p. 467) and is the primary unit of linguistic representation. The 'form' of a construction refers to and consists of its phonological/graphological and morphosyntactic features. 'Meaning' in CG comprises all of the conventionalized aspects of a given construction including its semantic, pragmatic, and discourse-functional properties. The need to include specific discourse-functional and pragmatic features among the defining properties of a construction is obvious if we consider expressions that seem to have exactly the same structure (e.g. Thank you! and See you!) but different semantic and pragmatic characteristics that sanction their use (Fried & Östman, 2004, p. 15). Constructions represent linguistic signs of differing size: from morphs to lexemes and multiword lexemes to the most general syntactic and semantic rules. The logical consequence of representing lexical items, larger linguistic patterns as well as regular syntactic and semantic phenomena as constructions, is the assertion that there is a continuum between lexical sign and syntactic constructions. Words and phrasal patterns are thus treated as equal contributors to building up complex linguistic expressions (p. 22).

2. *non-modularity* of form and meaning:

CG theorists argue that combining two or more forms usually does not result in a simple concatenation of the meanings the constituents have in isolation (Fried & Östman, 2004, p. 12), something that becomes obvious when one thinks of phraseologisms and other multiword lexemes, which present a problematic phenomenon for the traditional componential models of language and thus are in part responsible for the rise of CG (Croft, 2001, p. 15). Consequently, CG assumes that the form and the meaning of a construction do not form separate independent modules, but are inseparable and stand in a complex relationship to each other.

3. the *taxonomic organization* of constructions of one language.

The final tenet of CG is the assertion that the constructions of a given language do not simply form an irregular list of all patterns possible in that language. Instead, they reflect the linguistic conventions that the speakers of the language know and form a "structured inventory" of conventions (Langacker, 1987, pp. 63-76). The overall network organization of constructions in a language is best described as a taxonomic hierarchy of overlapping patterns, which are related through structure-shared properties.

Construction grammar has been applied successfully in first language acquisition research (Tomasello, 2003, 2007) and attempts have been made to employ it for linguistic analyses in SLA (e.g., Diessel, 2007; Haberzettl,

---

<sup>3</sup> There is some relation to Frame Semantics (Fillmore, 1982) and to Head-Driven Phrase Structure Grammar (Pollard & Sag, 1994; Sag, Wasow, & Bender, 2003). See for instance Riehemann (2001).

2007), but how applicable is it to CALL in general and ICALL in particular?

## *Construction Grammar in ICALL*

Let us return to Matthews' (1993) "criteria of adequacy" (p. 9): linguistic perspicuity, computational effectiveness, and acquisitional perspicuity and apply them to CG in ICALL.

### *Linguistic Perspicuity*

It should be evident from our sketch of CG in the previous section that it "offers a fruitful and insightful approach to analyzing language through a single conceptual tool—the notion of construction (Fried & Östman, 2004, p. 76). In itself, the concept of grammatical construction is "broad enough to represent every morphological or syntactic arguments/criteria/tests for identifying any syntactic category" and can be applied "to any grammatical structure, including both its form and its meaning" (Croft, 2001, p. 17). Kay (1997) describes CG—in the context of what we referred to earlier as the Berkeley School of CG—as "a non-modular, generative, non-derivational, monostratal, unification-based grammatical approach" (p. 123) which is committed to full coverage<sup>4</sup> and does not distinguish a priori between core and peripheral linguistic phenomena (pp. 126-129). Kay (2001), using the example of argument structure constructions, shows some advantages of a monotonic, purely unification-based approach to CG in comparison to approaches borrowed from Cognitive Linguistics (Goldberg, 1995). Of course, with the view to a possible computational implementation in an ICALL system, we need to rely on a grammatical formalism and thus orientate our understanding of CG largely on the Berkeley School. Figure 1 illustrates one way of formally representing any construction. Each box—form, meaning, and gradience<sup>5</sup>—stands for an attribute-value matrix or a tree of linguistic features and their respective values. A complete sub-tree, morphology, which is situated immediately under form is provided in Figure 3. Only the final nodes of the tree would be 'filled' with the appropriate values after an analysis of a construction. Each final node, with the exception of lemma which has a value of type string, will be given one of four possible values: uninstantiated (the default), unknown, true, false. In other words, the construction can be used as a 'container' for linguistic features which were uncovered during a computational textual analysis.

### *Computational Effectiveness*

Overall, one can state that the computational implementation of CG is still in its infancy, very few projects have been documented in the literature. However, Kay (1997, pp. 123-131, 2002) shows that it is possible to conceive of a formal architecture for CG, although he does not report a computational implementation. Then again, two major implementations of CG have been documented and both are used to conduct experiments on the "self-organization and evolution of language" (Steels, 2005, p. 213). The first CG formalism suitable for computational implementation is Embodied Construction Grammar (Bergen & Chang, 2002; Chang, Feldman, Porzel, & Sanders, 2002; Chang & Maia, 2001). Bryant (2004) refers to Embodied CG and explains a constructional

---

<sup>4</sup> It is with reference to full coverage that he wants the adjective 'generative' to be understood and not as an informal reference to "the family of GB [Government and Binding] related approaches" (Kay, 1997, p. 124).

<sup>5</sup> The concept of gradience in our view of construction grammar analysis of learner texts will be briefly outlined below.

analyzer. Embodied CG maps the form pole of a construction to its meaning pole by determining a best-fit (abstract) construction through unification and by relying on “four conceptual primitives that . . . suffice for building scalable language understanding systems: schemas, constructions, spaces and maps.” (Chang et al., 2002, p. 1). Schemas are conceptualizations of semantics, constructions are the main linguistic unit, maps capture cross-domain mappings of different entities (e.g., a doctor referring to a patient as ‘the gall bladder’), and spaces refer to the speaker’s/listener’s representation of the conceptual domain of the current discourse. In early work, Embodied CG is described as scalable, but not efficient (Chang & Maia, 2001). Given that the system has to be able to understand known constructions, interpret and learn new ones and generate conversational replies, the lack of efficiency arises because there are always a number of construction alternatives to process. Bryant (2004) points out that an Embodied CG analysis is possible if the constructional analyzer supports unification, multiple concurrent analyses, the flexible form relations and semantic expressiveness found in Embodied CG. Neither later papers about Embodied CG nor the next implementation note any lack of efficiency.

The second CG formalism—Fluid CG (Steels & De Beule, 2006)—is being developed and implemented at the SONY Computer Science Laboratory in Paris and in the Artificial Intelligence Laboratory at the Vrije Universiteit in Brussels. Steels and De Beule (2006) describe the formalism of Fluid CG and focus on the unification and merging algorithms. For the latter they introduce the J-operator (De Beule & Steels, 2005), a binary operator which enables the system after unifying two constructions, e.g., a determiner and a noun, to merge them to a new, more complex construction, e.g., a noun phrase.

The proposed formal architecture for CG and the two computational implementations are certainly a good starting position for development of a CG formalism in ICALL. It is our hypothesis that employing this grammatical formalism—with its exclusive reliance on the construction as the unit of analysis and unification to understand more or less grammatically and lexically specified constructions—will help us to avoid limitations of modeling methods which either rely on error anticipation or on constraint relaxation. A CG-based ICALL system ought to be able to identify well-formed and ill-formed constructions without attempting to anticipate either. In other words, the problem of a large search space in ICALL—learner texts with their huge variety of well-formed and ill-formed<sup>6</sup> utterances have to be described using preferably one and only one of enormously many, necessarily preconceived (mal-)rules or by selecting one of the many parses generated by relaxing the constraints<sup>7</sup> of the grammar—can be eliminated if the CG-based parser is written in such a way that it can learn constructions it has not yet encountered, similar to both Embodied and Fluid CG. This machine learning approach parallels our understanding of (second) language acquisition as an experienced-based learning process.

### *Acquisitional Perspicuity*

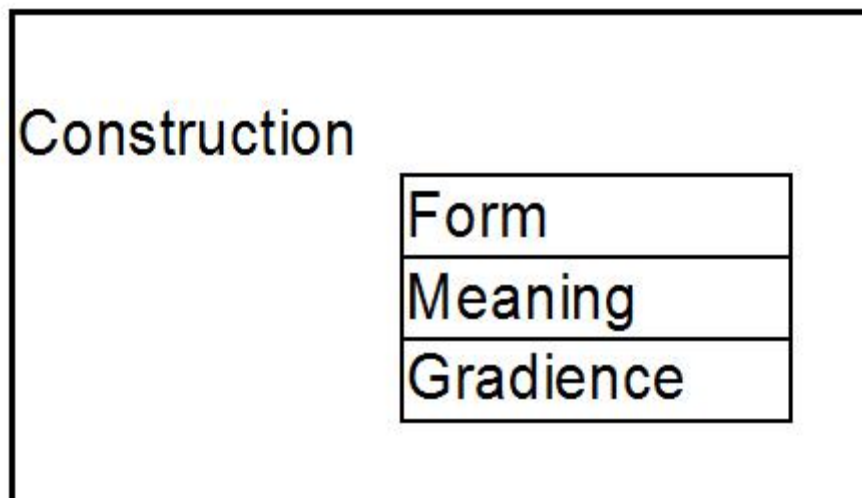
It is no coincidence that both implementations described above are situated in the domain of language acquisition. As already mentioned, CG has been applied successfully in first language acquisition research (Tomasello, 2003, 2007). Tomasello (2003) was able to show that young children acquire their first language by

---

<sup>6</sup> Often such utterances contain multiple errors.

<sup>7</sup> For an overview of different approaches to parsing ill-formed input in ICALL systems see Heift and Schulze (2007, pp. 34-43)

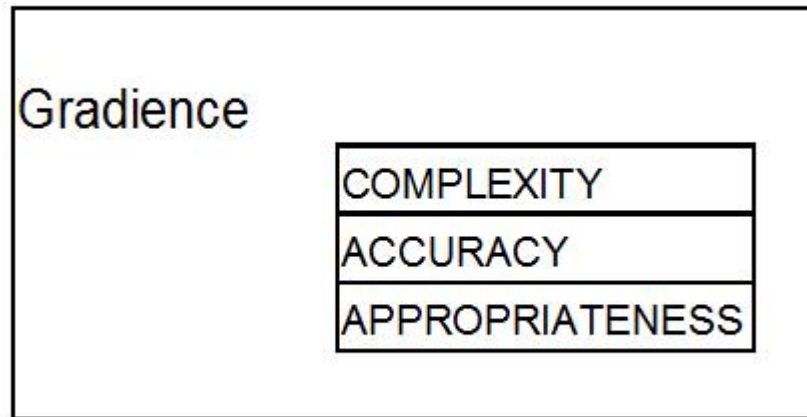
repeating holophrases first—short constructions which consist of fixed lexical material and have to be used with the same meaning and context. Later they manipulate acquired constructions and produce item-based constructions, in which clearly defined parts of the construction are substituted with other suitable items, to then arrive at the level of abstract constructions, which are (lexically) underspecified, complex linguistic signs and which could be compared to augmented phrase structure rules, e.g., the transitive-verb argument structure construction. Similar to Tomasello in L1, Haberzettl (2007) discusses “the emergence of constructions from prefabricated formulas” (p. 55). In general, the usage-based perspective of CG, its simultaneous consideration of complex forms and semantic, pragmatic and discourse-functional contexts as well as its adequacy for a dynamic, emergentist view of (second) language acquisition<sup>8</sup> (de Bot, Lowie, & Verspoor, 2005, 2007; Ellis & Larsen-Freeman, 2006; Larsen-Freeman, 1997, 2000, 2003, 2006) make it an ideal tool to map the search space of ICALL systems and to model individual language learning processes (Schulze, 2007). Since the grammatical framework does neither have to anticipate mal-rules (or positive transfer) from the learners L1, nor does it have to relax constraints of the learners L2, the CG framework can be used to capture the entire interlanguage continuum (Gass & Selinker, 2001; Selinker, 1974, 1992)—a characteristic which makes CG a suitable tool for SLA analysis and research. This can be done because a construction can be described formally by recording its features in an attribute-value matrix. For the analysis of texts produced by language learners, we need to complement the form-meaning mapping in the construction with information about complexity, accuracy, appropriateness which are subsumed under gradience (Bolinger, 1961) (see Figure 1 and Figure 2).



*Figure 1: Construction*

---

<sup>8</sup> Particularly the concept of grammar as a complex, dynamic (emergent) system in Fluid CG (Steels, 2005).



*Figure 2: Gradience*

The feature bundles—each represented by a box in the figures above—get filled during an interactive process of computational textual analysis. Our understanding of complexity, accuracy, and appropriateness as gradient feature bundles has the consequence that we do not make any binary decisions such as correct vs. incorrect or complex vs. simple or appropriate vs. inappropriate, instead we determine the probability of a particular utterance being of a particular complexity and accuracy level.<sup>9</sup>

This also means CG would facilitate the widening of the focus ICALL traditionally had (and to a some extent still has) on error analysis and corrective feedback. It remains to be seen whether or not CG frameworks can support error diagnosis and treatment similar to other formalisms that were employed in ICALL, but they will certainly be able to address problems such as multiword lexemes (Tschichold, 1999, 2000) on which ICALL researchers did not concentrate although they play an important role in second language learning. We are currently investigating the link between grammaticality judgments for selected constructions in learner texts and frequency of these constructions in certain small corpora (this approach is based on a study by Kempen & Harbusch, 2005), in order to address the problem with a CG formalism which might ‘learn’ well-formed and ill-formed new constructions equally ‘well’ and would therefore fail to provide error-contingent feedback. However, (negative) feedback on errors is not the only type of information learners want and need, CG frameworks with the rich information on form and meaning can provide, for instance, diagnostic feedback and information on the complexity of texts produced by learners. It is the latter in particular which is at the focus of our attention, we develop measures of learner text complexity to be able to better model language learning processes and to improve the individualization of ICALL systems.

### ***The Mocha Project***

Our work in this area is mainly done for the Mocha project.<sup>10</sup> We analyze longitudinal learner data—mainly

<sup>9</sup> For a discussion of our computational of complexity and accuracy see Schulze and Wood (2008).

<sup>10</sup> Principal investigator: Mathias Schulze; Co-investigator: Trude Heift; The research is supported by the Social Sciences and Humanities Research Council (SSHRC) of Canada; grant number 410-2007-2549.

foreign-language texts—from students at two Canadian universities and work on the development of student models, which are based on the analysis of learning outcomes over time. Careful plotting of language learning events such as the mastery of a particular abstract construction, in time series and phase space diagrams will provide us with the information required for our empirical investigation. Our analysis is solidly grounded in an integrative approach to second language acquisition (SLA) (de Bot et al., 2005, 2007; Ellis & Larsen-Freeman, 2006; Larsen-Freeman, 1997, 2000, 2003, 2006) in general and to student modeling in ICALL in particular (Schulze, 2007). Like Larsen-Freeman, de Bot, and others we view second language acquisition as a complex and dynamic system. Complexity here refers to the multitude of variables which affect the system, i.e. learning processes. These variables can be dependent on one another or interact in other ways, and thus change and are being changed in the process of learning a language. This change is the main reason why the system is described as dynamic. We explore and use methods, algorithms and mathematical models that have been developed in Dynamic Systems Theory (DST) (Hirsch, Smale, & Devaney, 2004; Luenberger, 1979) and Chaos Theory (Gleick, 1987; Lorenz, 1993; Williams, 1997) and which assist us in interpreting the plotted data. Our improved understanding of the complex and dynamic language learning processes will enable us to conceive, design, implement, and test learner models with our existing language learning software.

We are currently investigating the possibility of measuring the discourse complexity of learner texts and plotting these over time to get a good approximation of the individual learning process. In order to measure the complexity, we conduct a construction grammar analysis of each learner text and estimate the level of abstraction for each construction by calculating the frequency of the variants of that construction which the student produced before. Basically, we are attempting to get an approximation of the entropy in the text—how little or often constructions get repeated in text over time and how significantly one construction differs from another—with respect to the construction in question because we can then assume that text entropy and text complexity are proportional. Different instantiations of abstract constructions clearly result in a higher text entropy and are assumed to be an indication of a higher level of the complexity of the learner text. At the other end of the continuum, holophrase constructions which are identical to input material the student was likely to have seen are assumed to have a very low level of discourse complexity. Traditionally, this kind of complexity would have been described informally as the range of vocabulary and the range of grammatical constructions. The different complexity levels are then plotted in a phase space. This graph will give us some indication of what constructions learners used at different times in their language learning process and how they varied in discourse complexity relative to one another.

## ***Conclusion***

Construction Grammar is a grammatical framework which ought to be explored in ICALL. Its usage-based perspective and its emphasis on emerging language structures meshes well with current thinking about an integrative view of SLA (Larsen-Freeman & Cameron, 2008). The formalisms which have been developed thus far hold promise for a successful implementation in an ICALL system. CG formalisms facilitate the integration of learning algorithms for new, as yet unencountered constructions, which can be employed to address the problem of the large space in ICALL. Overall, Construction Grammar meets the three “criteria of adequacy” (Matthews,

1993, p. 9), its explanatory power in both (Formal) Linguistics and Second Language Acquisition is great and it is computationally effective.

## References

- Abeillé, A. (1992). The Lexicalised Tree Adjoining Grammar for French and Its Relevance to Language Teaching. In M. L. Swartz & M. Yazdani (Eds.), *Intelligent Tutoring Systems for Foreign Language Learning. The Bridge to International Communication* (pp. 65-88). Berlin: Springer Verlag.
- Bergen, B. K., & Chang, N. C. (2002). *Embodied Construction Grammar in Simulation-Based Language Understanding. Technical report 02-004*. Berkeley, CA: International Computer Science Institute.
- Bolinger, D. L. M. (1961). *Generality, Gradience, and the All-or-none*. 's-Gravenhage: Mouton.
- Brocklebank, C. P. (1998). *An Experiment in Developing a Prototype Intelligent Teaching System from a Parser Written in Prolog*. Unpublished MPhil Thesis, UMIST, Manchester.
- Bryant, J. (2004). Scalable Construction-Based Parsing and Semantic Analysis. In R. Porzel (Ed.), *HLT-NAACL 2004 Workshop: 2nd Workshop on Scalable Natural Language Understanding* (pp. 33-40). Boston: Association for Computational Linguistics.
- Bybee, J. (2006). From Usage to Grammar: The Mind's Response to Repetition. *Language*, 82(4), 711-733.
- Chang, N. C., Feldman, J., Porzel, R., & Sanders, K. (2002). *Scaling Cognitive Linguistics: Formalisms for Language Understanding*. Paper presented at the 1st International Workshop on Scalable Natural Language Understanding. from <http://www.icsi.berkeley.edu/NTL/papers/scaling.pdf>
- Chang, N. C., & Maia, T. V. (2001). *Grounded Learning of Grammatical Constructions*. Paper presented at the AAAI Spring Symposium on Learning Grounded Representations.
- Chanier, T., Pengelly, M., Twidale, M., & Self, J. A. (1992). Conceptual Modeling in Error Analysis in Computer-Assisted Language Learning Systems. In M. L. Swartz & M. Yazdani (Eds.), *Intelligent Tutoring Systems for Foreign Language Learning: The Bridge to International Communication* (pp. 125-150). New York: Springer Verlag.
- Chen, L., & Kurtz, B. L. (1989). XTRA-TE: Using Natural Language Processing Software to Develop an ITS for Language Learning. In D. Bierman, J. Breuker & J. Sandberg (Eds.), *Artificial Intelligence and Education. Proceedings of the 4th International Conference on AI and Education* (pp. 54-63). Amsterdam: IOS.
- Chomsky, N. (1986). *Knowledge of Language: Its Nature, Origin, and Use*. New York: Praeger.
- Croft, W. (2001). *Radical Construction Grammar. Syntactic theory in typological perspective*. Oxford: Oxford University Press.
- De Beule, J., & Steels, L. (2005). Hierarchy in Fluid Construction Grammar. In U. Furbach (Ed.), *Proceedings of the 28th Annual German Conference on AI, KI 2005* (pp. 1-15). Berlin: Springer.
- de Bot, K., Lowie, W., & Verspoor, M. (2005). Dynamic Systems Theory and Applied Linguistics: The Ultimate "so what"? *International Journal of Applied Linguistics*, 15(1), 116-118.
- de Bot, K., Lowie, W., & Verspoor, M. (2007). A Dynamic Systems Theory Approach to Second Language Acquisition. *Bilingualism: Language and Cognition*, 10(1), 7-21.
- Diessel, H. (2007). A Construction-Based Analysis of the Acquisition of East Asian Relative Clauses. *Studies in Second Language Acquisition*, 2007, 29, 2, June, 29(2), 311-320.
- Ellis, N. C., & Larsen-Freeman, D. (2006). Language Emergence: Implications for Applied Linguistics- Introduction to the Special Issue. *Applied Linguistics*, 2006, 27, 4, Dec, 27(4), 558-589.
- Feurman, K., Marshall, C., Newman, D., & Rypa, M. (1987). The CALLE Project. *CALICO Journal*, 4(2), 25-34.
- Fillmore, C. J. (1982). Frame semantics. In The Linguistic Society of Korea (Ed.), *Linguistics in the Morning Calm* (pp. 111-137). Seoul: Hanshin Publ. Co.
- Fischer, K., & Stefanowitsch, A. (2007). Konstruktionsgrammatik: Ein Überblick. In K. Fischer & A. Stefanowitsch (Eds.), *Konstruktionsgrammatik. Von der Anwendung zur Theorie* (pp. 3-17). Tübingen: Stauffenberg Verlag.
- Fried, M., & Östman, J.-O. (2004). Construction Grammar: a thumbnail sketch. *Construction Grammar in a cross-language perspective*, pp. 11-86,
- Fum, D., Pani, B., & Tasso, C. (1992). Naive vs. Formal Grammars: A Case for Integration in the Design of a Foreign Language Tutor. In M. L. Swartz & M. Yazdani (Eds.), *Intelligent Tutoring Systems for Foreign Language Learning. The Bridge to International Communication* (pp. 51-64). Berlin: Springer Verlag.
- Gass, S. M., & Selinker, L. (2001). *Second Language Acquisition. An Introductory Course*. Mahwah, NJ: Lawrence Erlbaum Associates.

- Gleick, J. (1987). *Chaos: Making a New Science*. New York, N.Y.: Viking.
- Goldberg, A. (1995). *Constructions. A Construction Grammar approach to argument structure*. Chicago: University of Chicago Press.
- Goldberg, A. (2006). *Constructions at Work: the nature of generalization in language*: Oxford University Press.
- Haberzettl, S. (2007). Konstruktionen im Zweitspracherwerb. In K. Fischer & A. Stefanowitsch (Eds.), *Konstruktionsgrammatik. Von der Anwendung zur Theorie* (pp. 55-77). Tübingen: Stauffenberg Verlag.
- Hamel, M.-J. (1996). NLP Tools in CALL for Error Diagnosis. *Revue de l'ACLA/Journal of the CAAL*, 18(2), 125-141.
- Handke, J. (1992). WIZDOM: A Multiple-Purpose Language Tutoring System Based on AI Techniques. In M. L. Swartz & M. Yazdani (Eds.), *Intelligent Tutoring Systems for Foreign Language Learning. The Bridge to International Communication* (pp. 293-306). Berlin: Springer Verlag.
- Heift, T. (1998). *Designed Intelligence: A Language Teacher Model*. Unpublished PhD Thesis, Simon Fraser University, Burnaby.
- Heift, T., & Nicholson, D. (2000). Theoretical and Practical Considerations for Web-based intelligent Language Tutoring Systems. In G. Gauthier, C. Frasson & K. VanLehn (Eds.), *Intelligent Tutoring Systems. 5th International Conference, IITS 2000, Montréal, Canada, June 19-23, 2000. Proceedings* (pp. 354-362). Montreal: Springer Verlag.
- Heift, T., & Schulze, M. (2007). *Errors and Intelligence in CALL. Parsers and Pedagogues*. New York: Routledge.
- Heinecke, J., Kunze, J., Menzel, W., & Schröder, I. (1998). Eliminative Parsing with Graded Constraints. In *Proceedings of Coling-ACL'98* (pp. 526-530).
- Hirsch, M. W., Smale, S., & Devaney, R. L. (2004). *Differential equations, dynamical systems, and an introduction to chaos* (2nd ed.). Amsterdam ; Boston ; London: Elsevier Academic Press.
- Kay, P. (1995). Construction Grammar. *Handbook of Pragmatics. Manual*, pp. 171-177,
- Kay, P. (1997). *Words and the Grammar of Context*. Stanford, CA: CSLI Publications.
- Kay, P. (2001). Argument Structure Constructions and the Argument-Adjunct Distinction. In M. Fried & H. C. Boas (Eds.), *Grammatical Constructions: Back to the Roots* (Eds.), Amsterdam: John Benjamins.
- Kay, P. (2002). An Informal Sketch of a Formal Architecture for Construction Grammar. *Grammars*, 5(1), 1-19.
- Kay, P., & Fillmore, C. J. (1999). Grammatical constructions and linguistic generalizations: The 'What's X doing Y?' construction. *Language*(75/1), 1-33.
- Kempen, G., & Harbusch, K. (2005). The Relationship between Grammaticality Ratings and Corpus Frequencies: A Case Study into Word Order Variability in the Midfield of German Clauses. In S. Kepser & M. Reis (Eds.), *Linguistic Evidence — Empirical, Theoretical, and Computational Perspectives* (pp. 329-349). Berlin: Mouton De Gruyter.
- L'Haire, S., & Vandeventer Faltin, A. (2003). Error Diagnosis in the FreeText Project. *CALICO Journal*, 20(3), 481-496.
- Labrie, G., & Singh, L. P. S. (1991). Parsing, Error Diagnostics and Instruction in a French Tutor. *CALICO Journal*, 9, 9-25.
- Lakoff, G. (1987). *Women, Fire, and Dangerous Things. What Categories Reveal about the Mind*. Chicago: University of Chicago Press.
- Langacker, R. W. (1987). *Foundations of Cognitive Grammar. Volume 1. Theoretical Prerequisites*. Stanford, CA: Stanford University Press.
- Larsen-Freeman, D. (1997). Chaos/Complexity Science and Second Language Acquisition. *Applied Linguistics*, 18(2), 141-165.
- Larsen-Freeman, D. (2000). Second Language Acquisition and Applied Linguistics. *Annual Review of Applied Linguistics*, 20, 165-181.
- Larsen-Freeman, D. (2003). *Teaching language : from grammar to grammaring*. Southbank, Victoria: Thomson/Heinle.
- Larsen-Freeman, D. (2006). The Emergence of Complexity, Fluency, and Accuracy in the Oral and Written Production of Five Chinese Learners of English. *Applied Linguistics*, 2006, 27, 4, Dec, 27(4), 590-619.
- Larsen-Freeman, D., & Cameron, L. (2008). *Complex Systems and Applied Linguistics*. Oxford: Oxford University Press.
- Levin, L. S., Evans, D. A., & Gates, D. M. (1991). The ALICE System: A Workbench for Learning and Using Language. *CALICO Journal*, 9(1), 27-54.
- Lorenz, E. N. (1993). *The Essence of Chaos*. Seattle: University of Washington Press.
- Luenberger, D. G. (1979). *Introduction to Dynamic Systems : Theory, Models, and Applications*. New York; Toronto Wiley.
- Matthews, C. (1993). Grammar Frameworks in Intelligent CALL. *CALICO Journal*, 11(1), 5-27.
- Matthews, C. (1994). Intelligent Computer Assisted Language Learning as Cognitive Science: The Choice of

- Syntactic Frameworks for Language Tutoring. *Journal of Artificial Intelligence in Education*, 5(4), 533-556.
- Menzel, W., & Schröder, I. (1999). Error Diagnosis for Language Learning Systems. In M. Schulze, M.-J. Hamel & J. Thompson (Eds.), *Language Processing in CALL. ReCALL Special Publication (Proceedings of a One-Day Conference "Natural Language Processing in Computer-Assisted Language Learning" Held at UMIST, 9 May 1998, Organised by the Centre of Computational Linguistics, UMIST, in Association with Eurocall)* (pp. 20-30). Hull: CTICML.
- Östman, J.-O., & Fried, M. (2004). Historical and intellectual background of Construction grammar. *Construction Grammar in a cross-language perspective*, pp. 1-10,
- Pereira, F. C. N., & Warren, D. H. D. (1980). Definite Clause Grammars for Language Analysis - A Survey of the Formalism and a Comparison with Augmented Transition Networks. *Artificial Intelligence*, 13, 231-278.
- Pijls, F., Daelemans, W., & Kempen, G. (1987). Artificial Intelligence Tools for Grammar and Spelling Instruction. *Instructional Science*, 16, 319-336.
- Pollard, C. J., & Sag, I. A. (1994). *Head-Driven Phrase Structure Grammar*. Chicago: University Chicago Press.
- Riehemann, S. Z. (2001). *A Constructional Approach to Idioms and Word Formation*. Stanford University, Stanford.
- Sag, I. A., Wasow, T., & Bender, E. M. (2003). *Syntactic Theory: A Formal Introduction* (2nd ed.). Stanford, Calif.: Center for the Study of Language and Information.
- Sanders, R. H. (1991). Error Analysis in Purely Syntactic Parsing of Free Input. The Example of German. *CALICO Journal*, 9, 72-89.
- Schönefeld, D. (2006). Constructions. *Constructions*, SV 1(1-2006). Retrieved from [http://www.constructions-online.de/articles/specvol1/667/Schoenefeld\\_Constructions.pdf](http://www.constructions-online.de/articles/specvol1/667/Schoenefeld_Constructions.pdf)
- Schulze, M. (1999). From the Developer to the Learner: Computing Grammar - Learning Grammar. *ReCall*, 11(1), 117-124.
- Schulze, M. (2001). *Textana - Grammar and Grammar Checking in Parser-Based CALL*. Unpublished PhD Thesis, UMIST, Manchester.
- Schulze, M. (2007). *Modeling SLA Processes Using NLP*. Paper presented at the Towards Adaptive CALL: Natural Language Processing for Diagnostic Assessment., Ames, IA.
- Schulze, M., & Hamel, M.-J. (1998). Use and Re-Use of Syntactic Parsers in CALL. Towards Diagnosing Learner Errors. In *Proceedings of WorldCALL* (pp. 203-204). Melbourne: Melbourne University.
- Schulze, M., & Wood, P. (2008). *Computational Analysis of Task-Based Performance*. Paper presented at the AAAL 2008. from [http://germanicandslavic.uwaterloo.ca/~mschulze/papers/aaal\\_2008.pdf](http://germanicandslavic.uwaterloo.ca/~mschulze/papers/aaal_2008.pdf)
- Schwind, C. B. (1990). An Intelligent Language Tutoring System. *International Journal of Man-Machine Studies*, 33(5), 557-579.
- Selinker, L. (1974). Interlanguage. In J. C. Richards (Ed.), *Error Analysis : Perspectives on Second Language Acquisition* (pp. 31-54). London: Longman.
- Selinker, L. (1992). *Rediscovering Interlanguage*. London: Longman.
- Steels, L. (2005). The Emergence and Evolution of Linguistic Structure: From Lexical to Grammatical Communication Systems. *Connection Science*, 17(3-4), 213-230.
- Steels, L., & De Beule, J. (2006). Unify and Merge in Fluid Construction Grammar. In P. Vogt, Y. Sugita, E. Tuci & C. L. Nehaniv (Eds.), *Symbol Grounding and Beyond, Third International Workshop on the Emergence and Evolution of Linguistic Communication, EELC 2006, Rome, Italy, September 30-October 1, 2006, Proceedings* (pp. 197-223): Springer.
- Tomasello, M. (2003). *Constructing a language : a usage-based theory of language acquisition*. Cambridge, Mass. ; London: Harvard University Press.
- Tomasello, M. (2007). Konstruktionsgrammatik und früher Erstspracherwerb. In K. Fischer & A. Stefanowitsch (Eds.), *Konstruktionsgrammatik. Von der Anwendung zur Theorie* (pp. 19-37). Tübingen: Stauffenberg Verlag.
- Tschichold, C. (1999). Grammar Checking for CALL: Strategies for Improving Foreign Language Grammar Checkers. In K. Cameron (Ed.), *Computer-Assisted Language Learning* (pp. 203-222). Lisse: Swets & Zeitlinger.
- Tschichold, C. (2000). *Multi-Word Units in Natural Language Processing*. Hildesheim: Olms Verlag.
- Vandeventer, A. (2000). Research on NLP-Based CALL at the University of Geneva. *TELL & CALL*, 2/2000, 10-13.
- Vandeventer, A. (2001). Creating a Grammar Checker for CALL by Constraint Relaxation: A Feasibility Study. *ReCALL*, 13(1), 110-120.
- Vandeventer, A., & Hamel, M.-J. (2000). Reusing a Syntactic Generator for CALL Purposes. *ReCALL*, 12(1), 79-91.
- Vandeventer Faltn, A. (2003). Natural Language Processing Tools for Computer Assisted Language Learning.

*Linguistik online*, 17(5), 137-153.

- Weinberg, A., Garman, J., Martin, J., & Merlo, P. (1995). A Principle-Based Parser for Foreign Language Tutoring in German and Arabic. In V. M. Holland, J. D. Kaplan & M. R. Sams (Eds.), *Intelligent Language Tutors: Theory Shaping Technology* (pp. 23-44). Mahwah, NJ: Lawrence Erlbaum Associates.
- Weischedel, R. M., Voge, W. M., & James, M. (1978). An Artificial Intelligence Approach to Language Instruction. *Artificial Intelligence*, 10, 225-240.
- Wilks, Y., & Farwell, D. (1992). Building an Intelligent Second Language Tutoring System from Whatever Bits You Happen to Have Lying Around. In M. L. Swartz & M. Yazdani (Eds.), *Intelligent Tutoring Systems for Foreign Language Learning. The Bridge to International Communication* (pp. 263-274). Berlin: Springer Verlag.
- Williams, G. P. (1997). *Chaos Theory Tamed*. Washington, D.C.: Joseph Henry Press.
- Zähner, C. (1991). Word Grammars in CALL. In H. Savolainen & J. Telenius (Eds.), *Eurocall 1991. International Conference on Computer Assisted Language Learning. Proceedings* (pp. 291-295). Helsinki: Helsinki School of Economics and Business Administration.

MORPHOLOGY

SEPARABILITY	
LEMMA	
POS	
	VERB
	WEAK VERB
	STRONG VERB
	MIXED VERB
	PRAET VERB
	MAIN VERB
	AUXVERB
	MODAL VERB
	FINITE VERB
	PAST PARTICIPLE
	PRES PARTICIPLE
	INFINITIVE
	NOUN
	MAIN NOUN
	PROPER NOUN
	PERSONAL PRONOUN
	DEMONSTRATIVE PRONOUN
	INDEFINITE PRONOUN
	QUANTIFICATIVE PRONOUN
	W PRONOUN
	POSSESSIVE PRONOUN
	DEFINITE ARTICLE
	INDEFINITE ARTICLE
	CONNECTOR
	POSTPOSITION
	PREPOSITION
	SUBORDINATING CONJUNCTION
	COORDINATING CONJUNCTION
	DISCONTINUOUS CONJUNCTION
	ADJECTIVE
	PRENOUN ADJ
	ADCOPULA ADJ
	ADVERB ADJ
	ADVERB
	PARTICLE
	MAIN ADVERB
	INTERJECTION
	RESPONSIVE INTERJECTION
	MAIN INTERJECTION

GENUS

MASCULINE

FEMININE

NEUTER

NUMERUS

SINGULAR

PLURAL

KASUS

NOMINATIVE

ACCUSATIVE

DATIVE

GENITIVE

GENUSVERBI

ACTIVE

PASSIVE

MIDDLE

TEMPUS

PRESENT TENSE

PAST TENSE

COMPARISON

POSITIVE

COMPARATIVE

SUPERLATIVE

MORPH POSITION

PREFIX

ROOT

CIRCUMFIX

SUFFIX

FUGENELEMENT

STATUS

MORPH

SUBLEX

SONSTITUENT

SATURATED

*Figure 3: Morphology*