The 'Learner Corpus Research, Cognitive Linguistics and Second Language Acquisition' nexus: a SWOT analysis.

SABINE DE KNOP and FANNY MEUNIER

Abstract

The introductory chapter of this special issue on 'Learner Corpus Research, Cognitive Linguistics and Second Language Acquisition' addresses the strengths, weaknesses, opportunities and potential threats of using both learner corpora and Cognitive Linguistics to research second language acquisition. We also discuss some terminological issues related to the notion of second language acquisition. Finally, we present the various chapters included in the volume and explain how each of them concretely articulates the connections between the three disciplines under analysis.

Keywords: learner corpus research, Cognitive Linguistics, second language acquisition

1 Introduction

This special issue aims to bolster interest in the collaborative research opportunities that exist between three disciplines, viz. Learner Corpus Research, Cognitive Linguistics and Second Language Acquisition. In this introductory chapter, we will first carry out a SWOT analysis (i.e. identify the Strengths, Weaknesses, Opportunities and Threats) of this nexus. We will then present the various papers included in the volume and explain how each of them concretely addresses the connections between the three disciplines under analysis.

Several publications have focused on the links between two of the three research domains listed above. As far back as the eighties, Langacker (1987, 1991a, 1991b), one of the founding fathers of Cognitive Linguistics, advocated a usage-based approach to language and cognition. Yet, despite the clear focus on usage, many linguists who adopted a cognitive approach still worked on constructed examples and failed to see advantages of using corpora and the authentic¹ language data they included. In 1999 Schönefeld explicitly stressed the complementarity of assumptions made by Cognitive Linguistics (CgL) and Corpus Linguistics (CpL^2) . Since then, several publications have convincingly demonstrated how the two frameworks could be combined (see for instance Gilquin 2010, Gries 2006, 2010 and 2012, Gries & Stefanowitsch 2006, Gonzales-Marquez et al. 2007). Other researchers have stressed the connections between CpL and pedagogy (see Granger et al. 2002, Hunston 2002, Hyland et al. 2012, and O'Keeffe et al. 2007). The advantages of crossing CgL and language pedagogy or CgL and second language acquisition (SLA) have also been described in numerous publications: see De Knop & De Rycker 2008, De Knop et al. 2010 for CgL and language pedagogy, and Robinson & Ellis 2008, Littlemore & Juchem 2010, Littlemore 2009, Skehan 1998, and Tyler 2012 for CgL and SLA. However, up to now little has been offered in terms of research on the intersection of the three domains, viz. CgL, CpL and SLA (theoretical or applied). Two articles dealing with more applied issues of language at the crossroads of CgL, CpL and language pedagogy came out in 2008 - "Corpora, cognition and pedagogical grammars: An account of convergences and divergences" (Meunier) and "Bridges between Cognitive Linguistics and second language pedagogy: The case of corpora and their potential" (Matusumoto) - but the first paper was mainly theoretical and

programmatic, and the second discussed a small sample of linguistic phenomena, e.g. the complementation patterns of the verb *find* by Japanese-speaking learners of English. More indepth studies investigating the interface between CgL, CpL – and more precisely learner corpus research – and SLA are still missing. This special issue on *Learner Corpus Research, Cognitive Linguistics and Second Language Acquisition* includes papers addressing this interface.

Before moving on to the SWOT analysis, we would briefly like to comment on the term 'second language acquisition'. SLA is often used in the literature as an umbrella term which, in actual facts, encompasses many different realities. Sociolinguistic cues are often used to distinguish second language acquisition from foreign language learning. Put simply, when a target language is learned in an environment where the language is used for everyday communication purposes, then the term second language acquisition applies; 'foreign language learning' characterizes the learning in an environment where the target language is not used for everyday communication purposes. This is however way too simplistic as numerous other variables come into play when a language other than the L1 is learned (see for instance the six volumes edited by Ortega in 2010 for a detailed review of all the critical concepts in SLA). Learning is often associated with conscious and voluntary processes whilst acquisition is associated with automatic, unconscious, proceduralized processes. The passage from conscious to unconscious processes is however not always easy to assess. The term 'second' language is also often used to refer to any language other than the L1, which means it can actually be a second, third, fourth, etc. language. Individual variation also plays an important role in the learning of new target languages. Key variables include, among others: age of learning; instruction (a so-called 'second language' can be learned without or with a little/some/intensive explicit instruction); types of motivation that lead a person to learn/use a language; attitudes towards that language; language aptitudes (in the L1 and L2/3/4/etc.); affective factors; amount and type of input received (some people in an SLA context receive less input than some others in an intensive foreign language context); amount and type of interaction in the target language; amount and type of output produced; etc. Because of the complex interplay between all those variables some 'late' (i.e. who typically started learning a new language after puberty) foreign language learners may end up having a much higher target language proficiency than early second language learners who have spent long years in the target language environment.

As the aim of this paper (and volume) is not to solve the problems related to SLA terminology, various labels will be used by the authors. The articles included in the volume mainly focus on foreign language contexts, or what Kachru (1986) would define as varieties of English that belong to the expanding circle and which are not norm-providing, nor norm-developing, but rather norm-dependent. All the target languages analysed in this volume can also be said to have been acquired/learned in 'instructed second language acquisition' (ISLA) contexts, but various terms will be found. They include 'second language acquisition', 'learners of a second language' or 'foreign learners'. This is due to the fact that each author has its own terminological preferences; but another reason is that when quoting other authors, the terms used initially are repeated, hence contributing to the terminological variety.

Before presenting the various articles included in the volume in more detail, we will first list the strengths, weaknesses, opportunities and potential threats of using both learner corpora and cognitive linguistics to research second language acquisition.

2 A SWOT evaluation

2.1 Strengths

Similar evolutionary trends and shared interests can be found between CgL, CpL and SLA.

The first one is related to the **form-meaning** relationship. CgL postulates form-meaning pairs, which means that there is no sharp distinction between lexical units and syntactic/grammatical structures. Grammatical structures are meaningful and differences in grammar reflect differences in meaning. Consequently, the differences between meaning and grammar are more a matter of degree rather than of conceptual nature. For Langacker (2008) grammar is the inventory of all linguistic units, i.e. of "lexical items, formatives, grammatical constructions, sound patterns, etc. which fluent speakers learn as units" (2008: 25). These conventional units - also called symbolic units - result from generalisations made in a bottom-up way during language acquisition on the basis of actual instances of language use, also called 'usage events' (2008: 25). Grammatical and lexical units can be of any length and of varying complexity, and they only differ from one another in terms of structural complexity. From Langacker's cognitive perspective, grammar has to describe not only the meanings of grammatical structures but also the grammatical categories used to build the constructions such as noun, verb, subject, object, clause, etc. Langacker's interpretation of the lexis-grammar continuum appears from his integration of word classes into the grammar of a language. The CgL framework will offer the possibility to describe examples found in corpora not only as single words but in the larger frame of syntactic structures. With the assumption of the symbolic character of grammar, it becomes clear that all data found in corpora are meaningful and that they deserve an analysis and description.

The lexis-grammar continuum is also central in CpL. As Römer summarizes it (2009: 141),

[i]f there is one major finding of modern (computer) corpus linguistic research over the past 40 years, it is probably that language is highly patterned. [...] Corpus studies, based on large collections of authentic text from a range of different sources, have provided massive evidence for the interdependence of lexis and grammar (or vocabulary and syntax). They have demonstrated that two areas that have traditionally been kept apart, both in language pedagogy and in linguistic theory, are in fact inseparable.

This focus on the patterned/formulaic/phraseological nature of language production is also key in language acquisition, be it first or second language acquisition. Emergentist views of language see its acquisition (Ellis 2008) as bottom-up generalizations made across multiple exposures to language input; language construction emerges from structural (lexis/grammar) regularities and is no longer considered as being innate or biologically predetermined. This is in total agreement with CgL's usage-based approach, where the description of specific aspects of language takes place starting with concrete examples, privileging a bottom-up approach. Speakers process linguistic information from specific utterances. "[T]he C[g]L approach is based on careful observation of how linguistic units are manipulated by language users in contexts of language use" (Tyler 2012: 216). This careful observation of authentic linguistic units is also at the core of the CpL enterprise. As claimed by Deshors (this issue: p.x) "meaning and abstract linguistic patterns emerge from speakers' experience of actual speech events". Specific utterances are collected in corpora and CpL offers strong tools to describe linguistic phenomena³. According to Gries (2008: 411) "C[g]L is the framework which makes most use of the framework of C[p]L". The usage-based model advocated by CgL can indeed strongly benefit from the quantitative corpus-linguistic methods, tools and language resources that have been developed over the last 40 years in CpL. Gries (2010: 336) adds that CgL has developed "theories and models that allow us to move from the purely descriptive approach for which C(p)L [our addition] is often criticized to explanation, prediction, and the embedding into a larger context, or theory, or model".

From a CpL perspective, the variety of authentic materials, consisting of written or spoken texts (see Deshors this issue), produced by different speakers, and in several registers and from different text types (e.g. fiction, press, academic or business language) constitutes an ideal raw material for cognitive linguists who can substantiate their claims with the help of such usage-based data.

Another central issue inherent to usage-based descriptions of language is **frequency**. Successful description of language phenomena must take into account the statistical properties of linguistic units. Taylor (2010: 40) disagrees with the generative approach to linguistic knowledge, in which frequency is regarded as some kind of epiphenomenon, arguing instead that "knowing a language involves, crucially, knowing the relative frequency of its various elements." Gries (2008: 414) even claims that "[t]he most basic corpuslinguistic tool is the frequency list". Frequency analysis makes it possible to describe the degree of entrenchment of particular units or structures, and offers a platform to research - for instance, why a specific word is being used in a particular construction - thereby giving cognitive grammar a possibility to deduce and describe different levels of generalization. Recurrent patterns are well-entrenched patterns in a specific language. As pointed out by Matsumoto (2008: 129), "[e]ntrenchment is interrelated with input. Specifically, entrenchment can be identified by an adjustment of the connection weights and can be brought about by the occurrence of a specific pattern of activation which renders more likely the occurrence of the same or a similar pattern." (for more details see also Bley-Vroman 2002, Ellis 2002, and Gries 2008). According to Stefanowitsch (2011: 260), "[m]any research questions can be investigated simply on the basis of whether or not a particular feature or element occurs under a given condition". Here the question is not so much how often a particular element is found in a corpus but also whether a linguistic item occurs or not in corpus data and more specifically in particular constructions. Of course, the distinction in occurrence vs. non-occurrence is strongly linked to the notion of frequency. Corpora can also be exploited as "a source of counter-examples [...] for disproving existing hypotheses" (Stefanowitsch 2011: 260). Learner corpora are a particularly rich source of such counterexamples which can inform linguists, teachers, and applied linguists about mistakes, problems, and difficulties arising in the process of foreign/second language acquisition. Granger et al. (2002), Granger (2009) and Hasko and Meunier (2013) provide insights into how learner corpora can contribute to second language acquisition studies. Recent developments in the collection of longitudinal learner corpora also reinforce the connections between SLA and CpL as more work is now being devoted to the study of the dynamics of L2 acquisition.

Applied perspectives of the research carried out is another case in point. The close connection between CgL and CpL can often be exploited for SLA issues. To become communicatively proficient in a foreign language (FL), learners need, among other things, to internalize not only the lexical units and the grammar of the FL, but also the **conceptualizations** underlying these lexical units and their distributional properties. The CgL enterprise postulates a strong link between language and cognition (Dirven & Ruiz de Mendoza Ibanez 2010) as CgL is said to approach "language as an integrated part of human cognition which operates in interaction with and on the basis of the same principles as other cognitive faculties" (Dirven 2004: 1). This postulate determines the basic principles and the methodology used in CgL to describe language. According to Langacker (1987: 5) "Cognitive Grammar [...] equates meaning with conceptualization" which means that all linguistic meaning is conceptual in nature and that linguistic entities reflect underlying concepts. The CgL approach offers explanations that "draw on learners' everyday real world experience by tapping into an intuitive reservoir of knowledge that facilitates an understanding of the systematic relationships among the units of language" (Tyler 2012: 18). But according to

Littlemore (2009: 29) "...the categorization systems that we build up due to our L1 cause us to form habits that are hard to break when we encounter a different language with different categorization systems." One of the difficulties that foreign language learners face is to recognize the foreign structures used to express meaning and to approach language as the reflection of underlying concepts which can vary from their L1 (see also Lantolf 1999). A CgL-based approach to SLA is motivated in the sense that it looks at the different categorizations underlying the foreign language. We will deal with these challenges and the difficulties related to L2 learning in section 2 and 3 hereunder. To remedy the categorization difficulties CgL offers some conceptual tools which can be useful in SLA, e.g. semantic networks, prototypes, embodiment, conceptual metaphor, and conceptual metonymy. SLA issues find their motivation in the application of such conceptual tools. E.g. prototypes like ongoingness or futurates are central in Littré's (this issue) study about the use of the English simple present vs. present progressive by French-speaking learners. "Strict but unmotivated and semi-accurate grammar rules and lists of exceptions are replaced by schematic patterns with central tendencies and extended, to more peripheral (but usually motivated) exemplars". (Tyler 2012: 62). In her book on Applying Cognitive Linguistics to Second Language Learning and Teaching, Littlemore (2009) quotes a series of studies that have shown that "non-native speakers tend to avoid using metaphorical senses of words, preferring to stick to more literal uses" (2009: 48). Using radial categories and teaching the more central senses first, to later introduce the more peripheral senses seems to foster the learning of foreign structures (see also Littlemore 2009). This in turn enhances memory (Boers 2011: 231). The recognition of related senses shows the motivation of senses and fosters better learning of foreign structures (Boers et al. 2010).

The pedagogical applications of learner corpus research are also central to the field. Granger (2003: 542) claims that "[1]earner corpus research opens up exciting pedagogical perspectives in a wide range of areas of English language teaching (ELT) pedagogy: materials design, syllabus design, language testing, and classroom methodology." Matsumoto (2008: 129) describes in more detail the positive potential of learner corpora for language pedagogy: "(i) they can help to decide what features should be particularly emphasized in teaching or even lead to the introduction of so far neglected elements; (ii) results from learner corpus studies can give indications on how to teach certain features; and (iii) results on developmental sequences can help to determine in what order language features should be taught. In other words, the more direct and probably more important way is to use a learner corpus to identify what is particularly difficult for a certain group of learners. The more indirect and more problematic way is to derive insights about second language acquisition from learner corpus analyses and to draw implications for teaching and possibly textbook writing from these insights." This is also confirmed by Deshors (this issue: p.x) who advocates the study of "co-occurrence patterns as they provide useful linguistic units to study L2 written linguistic structure and to capture emergent non-native linguistic patterns." But to collect corpus data will not be enough for successful SLA. "In second language acquisition, the role of entrenchment has been widely accepted as one of the most decisive factors in acquiring a second language. It is thus significant to pay attention to individual occurrences of linguistic items and at the same time to collect numerical data about types of construction." (Matsumoto 2008: 129). Littlemore (2009: 55ff) sees a possibility to exploit data from general corpora for data-driven learning (Johns 1991 and 1994, quoted in Littlemore 2009: 56; also Granger & Tribble 1998): It "involves showing learners multiple exemplars of a target language item taken from a corpus of authentic language, and asking them to develop their own ideas [...] about the possible meanings of that item." In her study of the uses of may and can by different English learner populations (i.e., French and Chinese) in speech and writing, Deshors (this issue: p.x) also highlights the importance of presenting second language learners with a statistical information about the distribution of linguistic items in L2 which further offers a possibility to recognize the systematic variation in those items' distribution across registers. Presentation alone may however not prove sufficient and enough care must also be devoted to authentication and pedagogical relevance (see Braun 2005, Belz 2004, Belz and Vyatkina 2008). Meunier (2011: 468) explains that language learners need to interpret, analyse and understand those linguistic items in a personally meaningful way. In order to do so, teachers can focus on specific linguistic forms produced by the learners themselves, in the context of meaningful interactions in communicative tasks. Learners are then more likely to feel a sense of authentication and pedagogical relevance. A comparison of data from general corpora with data from learner corpora can bring some additional dimensions: "In second language pedagogy, a data-driven analysis on the basis of both a learner corpus and a native speaker corpus is essential to explain the concept of entrenchment." (Matsumoto 2008: 129). Granger's (1996) Contrastive Interlanguage Analysis has highlighted the importance of comparing a learner corpora of the same target language.

2.2 Weaknesses

Despite the many shared interests listed in section 2.1, smooth interactions between disciplines which can also exist independently⁴ are not always possible; such interactions are sometimes even not desirable. We saw in the previous section that all linguistic meaning is conceptual in nature and that linguistic entities reflect underlying concepts. One problem, however, is that different speech communities can conceptualize reality in various, and not always corresponding, ways and may, as a result of this, use different linguistic units to express a similar reality (see De Knop's study on the semantic use of posture verbs, this issue). This may constitute a difficulty for foreign language learners as they have to 'revisit' some of their often deeply entrenched conceptualizations.

Another problem that has to be faced is how these underlying conceptualizations can be accessed when using learner corpora. Learner corpus analysis can reveal the difficulties learners may encounter with the conceptualization differences in various speech communities, at various stages of proficiency or in the mastery (or lack thereof) of specific genres, but the reasons behind these difficulties cannot be accessed easily. Some authors therefore recommend the combination of corpus data with other data types: Gries & Wulff (2005 and 2009) combine experimental evidence with native corpus data to analyse constructions in L2, and Meunier & Littré (2013) combine learner corpus data and experimental evidence to track progress in the acquisition of tenses and aspects by French EFL learners.

Another difference between CgL and CpL was put forward by Schönefeld (1999: 165). She explains that whilst both CgL and CpL "require that postulations and hypotheses about language be rooted in authentic language material", they nevertheless "differ in the degree to which they actually stick to this requirement. [...] Cognitive linguists, though also drawing on language data, do so in a less principal way." The statistical apparatus used for validating the research results is also typically more complex in CpL than in CgL. Myles (2008) also comments on the slow uptake of corpus tools and methods in SLA studies. As stated earlier already, each discipline has its own research routines, core features and jargon and it is not always easy to find a perfect equilibrium between what can also be considered as three types of deeply entrenched research habits. As will be clearly apparent from the papers included in the present volume, the cursor can often subtly be moved towards one or the other end of the three sides of the triangle (more corpus than cognitive, more descriptive than explanatory, less statistics but more applications, etc.). In line with Gries (2008: 421), we could claim that 'the theoretical ideal' would be that detailed studies of the behavior of any

symbolic unit integrate information from a variety of methods. In our imperfect reality, however, guaranteeing a fair and equally balanced integration of several research paradigms is often utopian.

Gries (2008: 422) also draws our attention to possible difficulties related to the correct retrieval of data, be it for frequency lists, collocations or concordances. These difficulties can concern the proper delimitation and definition of single units, e.g. in cases of compound words written with or without hyphen, differences between English and American spelling like *colour* vs. *color*, but also simply word boundaries. Another difficulty is dependent on the amount of data available in learner corpora: what amount is necessary to declare learner corpora as being representative of a linguistic use by learners? In her study of the abstract uses of the German verb *sitzen* De Knop (this issue) pinpoints this challenge as the number of hits for the verb *sitzen* is very low in the German learner corpora Falko. Gilquin (this issue) also discusses this problem in her study on phrasal verbs.

A least weakness that we find important to mention is the lack of a large body of 'effect of instruction' studies for both corpus- and/or cognitively-informed teaching applications. Such studies do exist (see for instance Tyler 2012 for an excellent review of studies applying cognitive linguistics to the teaching of English modal verbs, prepositions and clause level constructions, all including 'effect of instruction' measures; or Meunier (2012: 117-118) for references to studies addressing the impact of corpus consultation practices in promoting L2 proficiency) but many more should urgently be carried out to support (hopefully positive) evidence-based changes in teaching practices.

2.3 Threats⁵

One of the main difficulties with which L2 researchers can be confronted lies in the possible clash between L1 and L2 categories as "L1 categories exert strong priming effects that are transferred into the L2" (Littlemore 2009: 31). Boers et al. (2010: 5) even claim that L1 categorization can "create obstacles" for the learning of L2. In the same sense, Littlemore (2009: 29) stresses that "[t]hings become even more difficult for language learners when a concept that is divided into two broad categories in their own language is divided into, say, three categories in the target language". This is illustrated in the study by De Knop (this issue) about the semantic uses of posture verbs. Whereas French will use one very general term for location, i.e. *être* 'to be', German has a whole range of posture verbs which have to be used when speaking of the location of persons or objects.

We saw earlier in this chapter that the usage-based approach advocated by CgL can best be met by the use of authentic materials offered by corpora. But the argument of authenticity must also be relativized as the use of authentic material for SLA issues and teaching methodology can sometimes be problematic. As Gries notes (2008: 425) "it is still not always clear whether the learner benefits more from the exposure to authentic examples". Examples found in corpora can be too long, too complex for learners, which makes them unsuitable as illustrations for the learning of L2. Sometimes it is more effective to work with natural but nonetheless constructed examples for the exemplification of linguistic structures typical for L2. Similarly, Hunston (2002: 192) refers to Widdowson (2000) and Cook (1998) who see a danger of "extreme attitude towards using corpora in language teaching". She summarizes their arguments as follows:

1. Language in a corpus must be recontextualized in a pedagogic setting to make it real for learners.

2. Frequency should not be the only factor in deciding what to teach, i.e. teachers should not accept corpus evidence uncritically.

3. Corpora comprise the language of native speakers, whereas many learners will never communicate with a native speaker and/or are not interested in native speaker norms.

4. Learners should not be forced to approach English via 'lexical chunks' only.

The same need for caution holds for cognitively-inspired pedagogical applications. In his discussion of the added value of a Cognitive Semantics approach for SLA, Boers (2011: 246) quotes another important dimension which is often neglected in SLA issues, i.e. the profile of the learner. Teachers of foreign languages know very well that a consideration of the learners' proficiency level is decisive for a successful learning process. According to Boers (2011: 247) it is especially "the question whether C[ognitive]S[emantics]-style instruction is equally suitable for learners at different levels of proficiency". This issue would certainly merit more research. The concepts underlying CgL can indeed sound rather complex (embodied meaning, situated communication and construal, conceptual metaphor, mental imagery and spatial scenes, schemas, etc.) and, beyond any consideration related to proficiency in their L2, they may also need a relatively good level of conceptual understanding from the part of the learners (and teachers alike) if we want such an approach to make sense and prove useful. Here again the direct application of cognitive concepts to teaching may be risky and ad hoc pedagogical mediation is in order.

2.4 Opportunities

In her concluding chapter Hunston (2002: 213) summarizes the advantages of the use of corpora in applied linguistics (see also Ife 2004, Leech 1998, and Nesselhauf 2004). She claims that corpora have made life simpler, and indeed, the access to large collections of linguistic data from various registers, diverse groups, is easy. She goes on claiming that corpora have made language investigations possible that were not possible before: e.g. "[t]he corpus can show the diversity of use, and the importance, of very frequent words, current dictionaries tend to include more detailed information than the old ones" (Hunston 2002: 97). Corpora offer a platform for the "study of variation between languages produced in different situations" (Hunston 2002: 157). Learner corpora make it possible to identify the difficulties encountered by learners.

Gries (2012: 48) describes a positive side-effect of the application of corpus methods outside of CpL:

C[p]L would benefit from applying corpus methods outside of C[p]L and its discourses proper because that would increase C[p]L's visibility in the field of linguistics as a whole and in particular with disciplines that have often independently arrived at similar findings or conclusions, but also because external validation would streamline corpus-linguistic research enterprises.

But corpora, Hunston (2002: 213) adds, have also made life more complex: "The acquisition of linguistic elements and structures [...] involves extensive frequency-based processing of actual linguistic input, processing that in turn involves pattern matching, bottom-up categorization and inferencing, and storing (of instances/exemplars and/or schemas, [...])."Large collections of data would be rather unfruitful if they were not structured and exploited with an appropriate linguistic model. This is where CgL comes in. Choosing for a usage-based approach, and being further a "surface-oriented approach" (Gries 2008: 410), CgL offers the possibility to categorize and classify the data collected in corpora according to well-defined principles, e.g. in radial networks, with a prototypical sense and some peripheral ones. For Robinson & Ellis (2008: 494):

Cognitive Linguistics describes the properties of language in very different ways than either structuralist or generative approaches. There are no deep "structures" and no formal "rules" that generate permissible "strings" which the lexicon fills out. The product of learning reveals cross-linguistic differences in how languages structure conceptual content for expression, and cognitive linguists describe these differences. But the processes which give rise to them are shared by all language learners.

The intersection of CpL and CgL offers the possibility to exploit the common denominator and to describe these processes, and it seems that the time is now ripe to seize the opportunities offered by such interaction. As O'Keeffe & McCarthy (2010: 7-9) put it:

C[orpus] L[inguistics] leads to insights beyond the realms of lexis or grammar by applying its techniques to other questions, some more easily answered by computational analysis than others. In areas as diverse as second language acquisition and media studies, CL can be applied as a research tool. [...] Much of the purely descriptive research conducted by corpus linguists into language use would be of immense value to language teachers and material designers if more widely disseminated. [...] This process of engagement between CL and pedagogical applied linguistics needs to be improved.

As for Tyler (2012: 61), she explains how the theoretical tenets of Cognitive Linguistics "serve as conceptual tools to refine and reinvigorate familiar strands of L2 research such as cross-linguistic influence (or contrastive analysis), the use of authentic materials and implicit versus explicit language learning".

The papers included in the present volume all seek to demonstrate the value of combining the CgL and CpL frameworks to foster our understanding of the processes at play in SLA and to subsequently reinvigorate teaching practices.

3 Applications

The special issue aims to examine the meeting points in the interface of CgL and CpL starting from concrete applications. The articles focus on a variety of applied linguistic issues: syntax and semantics (phrasal verbs), tense and aspect (present progressive vs. present simple), the use of modal verbs (*can* instead of *may*), and the semantic uses of the posture verb *sitzen*. Data from various native and learner corpora, from different registers (writing or speech), from different L1-learners (French, Chinese), and for the learning of various L2s (English, German) are analysed and presented below.

Deshors examines how grammatical contexts constrain learners' linguistic choices differently in speech and writing; how English learners with different linguistic backgrounds develop different variation patterns across the two registers and what those distributional differences suggest as to what motivates the different patterns. She contrasts the uses of *may* and *can* by French and Chinese English learner populations and adopts a multidimensional methodological approach to corpus annotation and logistic regression modeling. The results show that French and Chinese English learners distinguish spoken and written uses of *may* and *can* on the basis of five grammatical features (type and semantics of modalized lexical verbs, negation, sentence type and voice). Deshors's study also stresses the need to involve several linguistic levels (i.e., semantic, morphological and syntactic) into single L2 register analyses.

De Knop analyses conceptual tools for the description and the acquisition of the German posture verb *sitzen* ('to sit'). She shows that some of the existing more abstract or metaphorical semantic uses of the German posture verb *sitzen* are not present in learners' written productions. She concludes that these abstract uses are not actively present in the learner's inventory of the semantic uses of posture verbs and further develops some teaching strategies for the acquisition of the more abstract uses by French-speaking learners. These strategies are based on semantic networks and conceptual metaphor.

Littré also uses a combination of corpus and experimental data to assess French L1 EFL learners' use and understanding of the simple and continuous present. He first reports on the results from an interpretation task where participants were found to show greater acceptance of more prototypical uses of the tenses and aspect under analysis. Learners also exhibited higher certainty with these prompts. The second part of his study examines written learner corpus data. The results show that (upper-intermediate to advanced) students still make errors involving basic functions of the simple and progressive. He also discusses the pedagogical implications of his results.

Gilquin's paper investigates the use of phrasal verbs by French-speaking foreign learners of English, using spoken and written learner corpus data and comparing them against native corpus data. She adopts a constructional approach and distinguishes between three levels of analysis: superconstruction, structural patterns and lexically specified phrasal verbs. The results show that the difficulties that learners have with phrasal verbs are mainly situated at the level of the superconstruction. The detailed results provided in her analysis can also be used to enhance the teaching of phrasal verbs for foreign language learners.

Although the number of studies in this special issue is limited, we hope to present some avenues for research issues at the crossroads of the three research domains, namely CgL, CpL and SLA.

Notes

- 1 Various types of corpus data are sometimes considered as being semi-authentic (rather than fully authentic) as they were elicited in instructed contexts.
- 2 As both cognitive linguistics and corpus linguistics are typically abbreviated as CL by their proponents, we will use the CgL and CpL abbreviations in the rest of the paper to avoid confusion.
- 3 We will not discuss the question of whether CpL is a methodology or a proper linguistic domain with a status similar to CgL. For more details see Gries (2010 and 2012) and Stefanowitsch (2011).
- 4 All corpus studies do not need to adopt a cognitive perspective; SLA studies can use many other data types than corpora; and the cognitive enterprise can easily exist without using learner corpora.
- 5 Contrary to what is typically the case in SWOT analyses, we have decided to address 'threats' before 'opportunities' as we prefer to link the opportunities with the description of the various papers of the volume.

References

Belz, Julie A. 2004. Learner corpus analysis and the development of foreign language proficiency. System 32(4): 577-591.

- Belz, Julie A. & Nina Vyatkina. 2008. The Pedagogical Mediation of a Developmental Learner Corpus for Classroom-Based Language Instruction. *Language Learning & Technology*, v. 12, n. 3, 33-52.
- Bley-Vroman, Robert. 2002. Frequency in production, comprehension, and acquisition. *Studies in Second Language Acquisition*, 24, 209–213.
- Boers, Frank. 2011. Cognitive Semantic ways of teaching figurative phrases. *Annual Review* of Cognitive Linguistics 9(1): 227-261.
- Boers, Frank, Antoon De Rycker and Sabine De Knop (eds.). 2010. Fostering Language Teaching Efficiency through CL: Introduction. De Knop, Sabine, Frank Boers & Antoon De Rycker (eds.). *Fostering Language Teaching Efficiency through CL*, 1-26. Berlin; New York: Mouton de Gruyter.
- Braun, Sabine. 2005. From pedagogically relevant corpora to authentic language learning contents. *ReCALL*, v. 17, n. 1, 47-64.
- Cook, Guy. 1998. The uses of reality: A reply to Ronald Carter. ELT Journal 52, 57-64.
- De Knop, Sabine & Teun De Rycker (eds.). 2008. Cognitive Approaches to Pedagogical Grammar: A Volume in Honour of René Dirven. Berlin; New York: Mouton de Gruyter.
- De Knop, Sabine, Frank Boers & Antoon De Rycker (eds.). 2010. *Fostering Language Teaching Efficiency through CL*. Berlin; New York: Mouton de Gruyter.
- Dirven, René. 2004. Major strands in Cognitive Linguistics. Essen: LAUD, Series A, Nr. 634.
- Dirven, René & Francisco Ruiz de Mendoza Ibanez. 2010. Looking back at 30 years of Cognitive Linguistics. In Tabakowska, Elzbieta, Michal Choinski & Lukasz Wiraszka (eds.), *Cognitive Linguistics in Action: From Theory to Application and Back*, 11-70. Berlin: de Gruyter Mouton.
- Ellis, Nick. 2002. Frequency effects in language processing and acquisition. *Studies in Second Language Acquisition* 24(2): 143-188.
- Ellis, Nick. 2008. Usage-based and form-focused language acquisition: The associative learning of constructions. In Peter Robinson and Nick Ellis (eds.), *Handbook of Cognitive Linguistics and Second Language Acquisition*, 372-406. London: Routledge.
- Gilquin, Gaëtanelle. 2010. *Corpus, Cognition and Causative Constructions*. John Benjamins Publishing Company : Amsterdam & Philadelphia.
- González-Márquez, Monica, Mittelberg, Irene, Coulson Seana & Spivey, Michael (eds.). 2007. Methods in Cognitive Linguistics. Amsterdam: John Benjamins.
- Granger, Sylviane. 1996. From CA to CIA and back: An integrated approach to computerized bilingual and learner corpora. In Aijmer K., Altenberg B. and Johansson M. ed(s) *Languages in Contrast. Text-based cross-linguistic studies*, 37-51. Lund: Lund University Press.
- Granger, Sylviane. 2003. The International Corpus of Learner English: A new resource for foreign language learning and teaching acquisition research. *TESOL Quarterly*, 37(3): 538-546.
- Granger, Sylviane. 2009. The contribution of learner corpora to second language acquisition and foreign language teaching: A critical evaluation, In Karin Aijmer (ed.), *Corpora and Language Teaching* 13-32. Amsterdam & Philadelphia : Benjamins.
- Granger, Sylviane & Tribble Chris. 1998. Learner corpus data in the foreign language classroom: Form-focused instruction and data-driven learning. In Granger Sylviane (ed.), *Learner English on Computer*, 199-209. London & New York: Addison Wesley Longman.
- Granger, Sylviane, Joseph Hung, & Stephanie Petch-Tyson (eds.). 2002. Computer Learner Corpora, Second Language Acquisition and Foreign Language Teaching. Amsterdam; Philadelphia: John Benjamins.

- Gries, Stefan T. 2006. Corpus-based methods and cognitive semantics: the many meanings of *to run*. In Gries, Stefan T. & Anatol Stefanowitsch (eds.), *Corpora in Cognitive Linguistics: Corpus-based Approaches to Syntax and Lexis*, 57-99. Berlin & New York: Mouton de Gruyter.
- Gries, Stefan T. 2008. Corpus-based methods in analyses of second language acquisition data. In Robinson, Peter & Nick Ellis (eds.), Handbook of Cognitive Linguistics and Second Language Acquisition, 406-431. New York: Taylor & Francis.
- Gries, Stefan T. 2010. Corpus linguistics and theoretical linguistics: a love-hate relationship? Not necessarily. *International Journal of Corpus Linguistics* 15 (3). 327-343.
- Gries, Stefan Th. 2012. Corpus linguistics, theoretical linguistics, and cognitive/psycholinguistics: towards more and more fruitful exchanges. In Joybrato Mukherjee & Magnus Huber (eds.), *Corpus Linguistics and Variation in English: Theory and Description*, 41-63. Amsterdam: Rodopi.
- Gries, Stefan T. & Anatol Stefanowitsch (eds.). 2006. *Corpora in Cognitive Linguistics*. *Corpus-based Approaches to Syntax and Lexis*. Berlin; New York: Mouton de Gruyter.
- Gries, Stefan T. & Stefanie Wulff. 2005. Do foreign language learners also have constructions? Evidence from priming, sorting, and corpora. *Annual Review of Cognitive Linguistics* 3: 182-200.
- Gries, Stefan Th. & Stefanie Wulff. 2009. Psycholinguistic and corpus-linguistic evidence for L2 constructions. *Annual Review of Cognitive Linguistics* 7: 163-186.
- Hasko, Victoria & Fanny Meunier (eds.). 2013. *Capturing L2 Development Through Learner Corpus Analysis*, The Modern Language Journal, Volume 97, Supplement 2013, Issue S1.
- Hunston, Susan. 2002. Corpora in Applied Linguistics. Cambridge: Cambridge University Press.
- Hyland, Ken, Chau Meng Huat, & Michael Handford (eds.). 2012. Corpus Applications in Applied Linguistics. London: Continuum.
- Ife, Anne. 2004. The L2 learner corpus: Reviewing its potential for the early stages of learning. In Baynham Mike, Alice Deignan, and Bill White (eds), *Applied Linguistics at the Interface*, 91-103. London/Oakville: British Association for Applied Linguistics in association with Equinox.
- Johns, Tim F. 1991. Should you be persuaded: two examples of data-driven learning. In T.F. Johns and P. Kind (eds.), *Classroom Concordancing*, 1-13. Birmingham: English Language Research.
- Johns, Tim F. 1994. From printout to handout: Grammar and vocabulary teaching int the context of data-driven learning. In Terence Odlin (ed.), *Perspectives on Pedagogical Grammar*, 293-313. New York: Cambridge University Press.
- Kachru, Braj. 1986. *The Alchemy of English: the spread, functions and models of non-native Englishes*. Oxfrod: Pergamon Press.
- Langacker, Ron W. 1987. Foundations of Cognitive Grammar. Vol. 1: Theoretical Prerequisites. Stanford, Calif.: Stanford University Press.
- Langacker, Ron W. 1991a. Foundations of Cognitive Grammar. Vol. 2: Descriptive Application. Stanford, Calif.: Stanford University Press.
- Langacker, R. W. 1991b. *Concept, Image, and Symbol: The Cognitive Basis of Grammar*. Berlin; New York: Mouton de Gruyter.
- Langacker, Ron W. 2008. The relevance of Cognitive Grammar for language pedagogy. In De Knop, Sabine & Teun De Rycker (eds.). 2008. Cognitive Approaches to Pedagogical Grammar: A Volume in Honour of René Dirven, 7-35. Berlin; New York: Mouton de Gruyter.

- Lantolf, James P. 1999. Second culture acquisition: Cognitive considerations. In Eli Hinkel (ed.), *Culture in Second Language Teaching and Learning*, 28-46. Cambridge: Cambridge University Press
- Leech G. 1998. Learner corpora: what they are and what can be done with them. In Granger Sylviane (ed.), *Learner English on Computer*, xiv-xx. London; New York: Addison Wesley Longman.
- Littlemore, Jeannette. 2009. Applying Cognitive Linguistics to Second Language Learning and Teaching. London: Palgrave MacMillan.Littlemore, Jeannette & Constanze Juchem-Grundmann (eds.). 2010. Applied Cognitive Linguistics in Second Language Learning and Teaching. Special issue of AILA Review 23.
- Matsumoto, Noriko. 2008. Bridges between Cognitive Linguistics and second language pedagogy: The case of corpora and their potential. *SKY Journal of Linguistics* 21: 125–153.
- Meunier, Fanny. 2008. Corpora, cognition and pedagogical grammars: An account of convergences and divergences. In De Knop, Sabine & Teun De Rycker (eds.), *Cognitive Approaches to Pedagogical Grammar*. A Volume in Honour of René Dirven, 91-119. Berlin: Mouton de Gruyter.
- Meunier, Fanny. 2011. Corpus linguistics and second/foreign language learning: exploring multiple paths. *Revista Brasileira de Linguística Aplicada*, Vol. 11, no. 2, 459-477.
- Meunier, Fanny. 2012. Formulaic Language and Language Teaching. In : Annual Review of Applied Linguistics, no. 1 (2012), p. 111-129.
- Meunier, Fanny & Damien Littré. 2013. Tracking Learners'Progress: Adopting a Dual 'Corpus cum Experimental Data' Approach. *Modern Language Journal*, Vol. 97, no. S1, 61-76.
- Myles, Florence. 2008. Investigating learner language development with electronic longitudinal corpora: Theoretical and methodological issues. In L. Ortega & H. Byrnes (Eds.), The longitudinal study of advanced L2 capacities (pp. 58–72). New York/London: Routledge.
- Nesselhauf, Nadja. 2004. Learner corpora and their potential for language teaching. In Sinclair, John (ed.), *How to Use Corpora in Language Teaching*, 125–152. Amsterdam: John Benjamins.
- O'Keeffe, Anne and McCarthy, Micheak (eds.) 2010. *The Routledge Handbook of Corpus Linguistics*. London: Routledge.
- O'Keeffe, Anne, Michael McCarthy, & Ronald Carter. 2007. From Corpus to Classroom. Language use and Language Teaching. Cambridge: Cambridge University Press.
- Ortega, Lourdes (ed.) 2010. Second-Language Acquisition (Critical Concepts in Linguistics). London: Routledge.
- Robinson, Peter & Nick Ellis (eds.). 2008. *Handbook of Cognitive Linguistics and Second Language Acquisition*, New York: Taylor & Francis.
- Römer, Ute. 2009. The inseparability of lexis and grammar. Corpus linguistic perspectives. *Annual Review of Cognitive Linguistics* 7: 141–163.
- Schönefeld, Doris. 1999. Corpus linguistics and cognitivism. *International Journal of Corpus Linguistics* 4(1): 137-171.
- Skehan, Peter. 1998. A Cognitive Approach to Language Learning. Oxford: Oxford University Press.
- Stefanowitsch, Anatol. 2011. Cognitive linguistics meets the corpus. In Brdar, Mario, Stephan T. Gries, and Milena Zic Fuchs (eds.), *Cognitive Linguistics: Convergence and Expansion*, 257–290. Amsterdam; Philadelphia: John Benjamins.

- Taylor, John. 2010. Language in the Mind. In De Knop, Sabine, Frank Boers, and Antoon De Rycker (eds.), *Fostering Language Teaching Efficiency through CL*, 29-57. Berlin/New York: Mouton de Gruyter.
- Tyler, Andrea (2012). Cognitive Linguistics and Second Language Learning. New York; London: Routledge.
- Widdowson, Henry G. 2000. On the limitations of linguistics applied. *Applied Linguistics* 21, 3-25.