

1

**Cognition, corpora, and computing [Recurso electrónico] : triangulating research in usage-based language learning / Nick C. Ellis**

Este artículo se encuentra disponible en su edición electrónica. Su acceso electrónico es a través del enlace de 'Acceso al documento'.

References: p. 57-65

Usage-based approaches explore how we learn language from our experience of language. Related research thus involves the analysis of the usage from which learners learn and of learner usage as it develops. This program involves considerable data recording, transcription, and analysis, using a variety of corpus and computational techniques, many of them specially devised for learner language. This article surveys relevant developments across the psychology of learning, first and second language acquisition, psycholinguistics, corpus linguistics, and computational linguistics and identifies challenges and future priorities relating to the following issues: (1) analyzing the distributional characteristics of linguistic constructions and their meanings in large collections of language that are representative of the language that learners experience, (2) the longitudinal analysis of learner language, and (3) Natural Language Processing analyses of the dimensions of language complexity.

Language learning. -- 2017 (June), v. 67, n. 1, supp. 1, p. 40-65

1. Cognitive linguistics 2. Corpus linguistics 3. Language learning 4. Psycholinguistics 5. Usage-based approaches

2

**Collocations in corpus-based language learning research [Recurso electrónico] : identifying, comparing, and interpreting the evidence/ Dana Gablasova, Vaclav Brezina, and Tony McEnery**

Este artículo se encuentra disponible en su edición electrónica. Su acceso electrónico es a través del enlace de 'Acceso al documento'.

References: p. 175-179

This article focuses on the use of collocations in language learning research (LLR). Collocations, as units of formulaic language, are becoming prominent in our understanding of language learning and use; however, while the number of corpus-based LLR studies of collocations is growing, there is still a need for a deeper understanding of factors that play a role in establishing that two words in a corpus can be considered to be collocates. In this article we critically review both the application of measures used to identify collocability between words and the nature of the relationship between two collocates. Particular attention is paid to the comparison of collocability across different corpora representing different genres, registers, or modalities. Several issues involved in the interpretation of collocational patterns in the production of first language and second language users are also considered. Reflecting on the current practices in the field, further directions for collocation research are proposed.

Language learning. -- 2017 (June), v. 67, n. 1, supp. 1, p. 155-179

1. Association measures 2. Collocations 3. Corpus linguistics 4. Formulaic language 5. Second language acquisition

3

**Combining language corpora with experimental and computational approaches for language acquisition research [Recurso electrónico] / Padraic Monaghan, and Caroline F. Rowland**

Este artículo se encuentra disponible en su edición electrónica. Su acceso electrónico es a través del enlace de 'Acceso al documento'.

References: p. 33-39

Historically, first language acquisition research was a painstaking process of observation, requiring the laborious hand coding of children's linguistic productions, followed by the generation of abstract theoretical proposals for how the developmental process unfolds. Recently, the ability to collect large-scale corpora of children's language exposure has revolutionized the field. New techniques enable more precise measurements of children's actual language input, and these corpora constrain computational and cognitive theories of language

development, which can then generate predictions about learning behavior. We describe several instances where corpus, computational, and experimental work have been productively combined to uncover the first language acquisition process and the richness of multimodal properties of the environment, highlighting how these methods can be extended to address related issues in second language research. Finally, we outline some of the difficulties that can be encountered when applying multimethod approaches and show how these difficulties can be obviated.

Language learning. -- 2017 (June), v. 67, n. 1, supp. 1, p. 14-39

1. Computational modeling 2. Corpus analysis 3. First language acquisition 4. Multiple cues 5. Second language acquisition

---

#### 4

#### **Empirical learner language and the levels of the 'Common European Framework of Reference' [Recurso electrónico] / Katrin Wisniewski**

Este artículo se encuentra disponible en su edición electrónica. Su acceso electrónico es a través del enlace de 'Acceso al documento'.

References: p. 247-253

The Common European Framework of Reference (CEFR) is the most widespread reference tool for linking language tests, curricula, and national educational standards to levels of foreign language proficiency in Europe. In spite of this, little is known about how the CEFR levels (A1–C2) relate to empirical learner language(s). This article sums up recent trends to meet the need of empirical CEFR level research, where learner corpus-based analyses play an increasing role. A first focus of the article is on studies that aim at illustrating CEFR levels by analyzing rated learner texts ("criterial features"). Furthermore, research that tries to disentangle the empirical validity of the CEFR scales by operationalizing its descriptors is presented. Before concluding with an outline of the most urgent research needs, potentials, and boundaries of interdisciplinary work between the fields of language testing and assessment, second language acquisition, and learner corpus research are discussed.

Language learning. -- 2017 (June), v. 67, n. 1, supp. 1, p. 232-253

1. CEFR 2. CEFR scales 3. Criterial features 4. Learner corpora 5. Proficiency levels 6. Scale validity

---

#### 5

#### **Evidence and interpretation in language learning research [Recurso electrónico] : opportunities for collaboration with computational linguistics / Detmar Meurers, and Markus Dickinson**

Este artículo se encuentra disponible en su edición electrónica. Su acceso electrónico es a través del enlace de 'Acceso al documento'.

References: p. 88-95

This article discusses two types of opportunities for interdisciplinary collaboration between computational linguistics (CL) and language learning research. We target the connection between data and theory in second language (L2) research and highlight opportunities to (a) enrich the options for obtaining data and (b) support the identification and valid interpretation of relevant learner data. We first characterize options, limitations, and potential for obtaining rich data on learning: from Web-based intervention studies supporting the collection of experimentally controlled data to online workbooks facilitating large-scale, longitudinal corpus collection for a range of learning tasks and proficiency levels. We then turn to the question of how corpus data can systematically be used for L2 research, focusing on the central role that linguistic corpus annotation plays in that regard. We show that learner language poses particular challenges to human and CL analysis and requires more interdisciplinary discussion of analysis frameworks and advances in annotation schemes.

Language learning. -- 2017 (June), v. 67, n. 1, supp. 1, p. 66-95

1. Annotations schemes 2. Computational linguistics 3. Corpus annotation 4. ICALL 5. Learner corpora

6

**Exploring learner language through corpora [Recurso electrónico]: comparing and interpreting corpus frequency information / Dana Gablasova, Vaclav Brezina, and Tony McEney**

Este artículo se encuentra disponible en su edición electrónica. Su acceso electrónico es a través del enlace de 'Acceso al documento'.

References: p. 150-153

This article contributes to the debate about the appropriate use of corpus data in language learning research. It focuses on frequencies of linguistic features in language use and their comparison across corpora. The majority of corpus-based second language acquisition studies employ a comparative design in which either one or more second language (L2) corpora are compared to a first language (L1) production corpus or two or more L2 corpora are compared to each other. This article critically examines some of the central tenets of the comparative method related to the interspeaker variation in L1 and L2 use, the representativeness and comparability of corpus data, the interpretation of difference found between corpora and the appropriate use of statistics. Using and discussing a set of five L1 spoken English corpora and three L2 English corpora (two spoken and one written), we approach these areas empirically exploring different sources of variations and methodological options that corpus-based SLA studies offer.

Language learning. -- 2017 (June), v. 67, n. 1, supp. 1, p. 130-154

1. Comparative design 2. Comparing L1 and L2 speakers 3. Corpus frequency 4. Corpus-based SLA 5. Native speaker corpora 6. Nonnative speaker corpora

7

**Interdisciplinary research at the intersection of CALL, NLP, and SLA [Recurso electrónico] : methodological implications from an input enhancement project / Nicole Ziegler ... [et al.]**

Este artículo se encuentra disponible en su edición electrónica. Su acceso electrónico es a través del enlace de 'Acceso al documento'.

References: p. 227-231

Despite the promise of research conducted at the intersection of computer-assisted language learning (CALL), natural language processing, and second language acquisition, few studies have explored the potential benefits of using intelligent CALL systems to deepen our understanding of the process and products of second language (L2) learning. The strategic use of technology offers researchers novel methodological opportunities to examine how incremental changes in L2 development occur during treatment as well as how the longitudinal impacts of experimental interventions on L2 learning outcomes occur on a case-by-case basis. Drawing on the pilot results from a project examining the effects of automatic input enhancement on L2 learners' development, this article explores how the use of technology offers additional methodological and analytical choices for the investigation of the process and outcomes of L2 development, illustrating the opportunities to study what learners do during visually enhanced instructional activities.

Language learning. -- 2017 (June), v. 67, n. 1, supp. 1, p. 209-231

1. Input enhancement 2. Intelligent computer-assisted language learning 3. Natural language processing 4. Research methods 5. Second language acquisition

8

**Linguistic models, acquisition theories, and learner corpora [Recurso electrónico] : morphological productivity in SLA research exemplified by complex verbs in German / Anke Lüdeling, Hagen Hirschmann, and Anna Shadrova**

Este artículo se encuentra disponible en su edición electrónica. Su acceso electrónico es a través del enlace de 'Acceso al documento'.

References: p. 125-129

The present study analyzes morphological productivity for complex verbs in second language acquisition by analyzing a corpus of German as a Foreign Language (GFL). It shows that advanced learners of GFL use prefix and particle verbs relatively frequently and productively but less so than native speakers do and discusses these findings in the light of different linguistic models and acquisition theories. It argues that corpus data must be

evaluated against good models and that it is necessary to make the categorization decisions available as annotations.

Language learning. -- 2017 (June), v. 67, n. 1, supp. 1, p. 96-129

1. Annotations 2. Complex verbs 3. Learner corpora 4. Morphological productivity

---

9

**A shared platform for studying second language acquisition [Recurso electrónico] / Brian MacWhinney**

Este artículo se encuentra disponible en su edición electrónica. Su acceso electrónico es a través del enlace de 'Acceso al documento'.

References: p. 269-275

The study of second language acquisition (SLA) can benefit from the same process of datasharing that has proven effective in areas such as first language acquisition and aphasiology. Researchers can work together to construct a shared platform that combines data from spoken and written corpora, online tutors, and Web-based experimentation. Many of the methods and tools for building this platform are already available as a result of earlier work on corpus sharing in first language acquisition. By working together on a shared platform in a coordinated manner, researchers will be able to construct a rich new empirical basis for the study of SLA.

Language learning. -- 2017 (June), v. 67, n. 1, supp. 1, p. 254-275

1. Corpus linguistics 2. Data-sharing 3. Language learning 4. Multimedia 5. Spoken language 6. Tutorials

---

10

**Task effects on linguistic complexity and accuracy [Recurso electrónico] :a large-scale learner corpus analysis employing natural language processing techniques / Theodora Alexopoulou, Marije Michel, Akira Murakami, and Detmar Meurers**

Este artículo se encuentra disponible en su edición electrónica. Su acceso electrónico es a través del enlace de 'Acceso al documento'.

References: p. 203-207

Large-scale learner corpora collected from online language learning platforms, such as the EF-Cambridge Open Language Database (EFCAMDAT), provide opportunities to analyze learner data at an unprecedented scale. However, interpreting the learner language in such corpora requires a precise understanding of tasks: How does the prompt and input of a task and its functional requirements influence task-based linguistic performance? This question is vital for making large-scale task-based corpora fruitful for second language acquisition research. We explore the issue through an analysis of selected tasks in EFCAMDAT and the complexity and accuracy of the language they elicit.

Language learning. -- 2017 (June), v. 67, n. 1, supp. 1, p. 180-208

1. Accuracy 2. Complexity 3. Fluency (CAF) 4. Learner corpus 5. NLP 6. Task complexity 7. TBLT

---