

Review of / Ressenya de / Reseña de

Gudmestad, Aarnes; Edmonds, Amanda (Eds.) (2018). *Critical Reflections on Data in Second Language Acquisition*. John Benjamins

ISBN: 9789027201423 (hard copy); 9789027201430 (paperback); 9789027263551 (e-book);
DOI: <https://doi.org/10.1075/llt.51>

Jennifer THEWISSEN

University of Antwerp & Université Saint-Louis, Belgium

jennifer.thewissen@uantwerpen.be

<https://orcid.org/0000-0003-1298-4679>

Keywords: second language instruction, L2 data, methodology, transparency, open science

Català:

Paraules clau: ensenyament de segones llengües; dades de L2; metodologia; transparència, ciència oberta

Español:

Palabras clave: enseñanza de segundas lenguas; datos de L2; metodología; transparencia; ciencia abierta

This volume, edited by Aarnes Gudmestad and Amanda Edmonds, offers an in-depth critical reflection on a key methodological aspect of second language acquisition (SLA) research, namely “data”. The book is testimony to the growing awareness of the need for increased methodological rigor and transparency in SLA studies and, as such, is situated within what Byrnes (2013, p. 825) has termed the “methodological turn” in SLA. The editors explain that the compilation of the volume was prompted by the observation that each step in SLA research design (e.g. construct definitions, data collection, data coding, statistical analyses, qualitative analyses, etc.) involves crucial decisions that will impact the findings yielded about L2 development. These key questions, however, very much tend to be discussed and solved internally among individual research teams. The book is a call to move methodological considerations into the open so that SLA researchers can collectively benefit from each other’s expertise and solutions. The scholars who contributed to the volume have generously done so by sharing their own methodological experience in specialized SLA areas ranging from interlanguage pragmatics, phonology, proficiency, corpus linguistics, or concept-oriented analysis, to name but a few.



The volume is organized around three main sections, each showcasing a specific type of critical examination. Part 1 involves building a corpus of published studies in a specific domain to identify the core methodological issues that stand out from the collected literature. Part 2 is based on the recoding and reanalysis of published work to study the impact that methodological changes have on SLA findings. Part 3 zooms in on specific approaches to SLA (e.g., corpus linguistics, concept-oriented analyses). Each section is illustrated by means of two or three contributions, all of which are written as a critical reflection on methodological practices in the collection, coding and/or analysis of SLA data. The volume includes a total of eight chapters along with a conclusion by Emma Marsden and Luke Plonsky. The general volume structure is outlined below along with a brief description of the chapters in each section.

Part 1 is an analysis of methodological practices in specific fields of SLA based on systematic literature reviews of papers selected from high-quality journals. It includes three chapters.

- Chapter 1 was written by Naoko Taguchi. The author focuses on three data types which have been used to study interlanguage pragmatic development, namely (written and spoken) discourse completion tasks, role plays, and naturalistic recordings. Her literature review highlights that these data types (including the more naturalistic data) regularly consist of utterance-level analyses and are largely based on frequency counts of the specific linguistic forms used to express a given speech act. She calls for an increased research focus on pragmatic performance in interactional settings to illustrate how learners collaboratively co-construct a speech act and how this ability evolves over time.
- Chapter 2, by Shaofeng Li, provides a highly rigorous synthesis of the data collection methods used in oral corrective feedback (CF) studies. CF refers to “the responses to the errors that learners commit” (p. 34) in the production or comprehension of the L2. The author’s review of the literature points to heterogeneous research practices in three core areas: CF treatment, CF elicitation tasks, and measurement of CF effects. Importantly, this chapter systematically reminds the reader of the need to safeguard three key methodological concepts in study design, namely internal, external/ecological, and construct validity in CF studies. Internal validity refers to the degree of confidence that the causal relationship being tested is not influenced by other factors or variables; external validity indicates the extent to which results can be generalized to other learner groups, while construct validity is understood as the extent to which a study measures the construct it is supposed to measure.
- Chapter 3 was written by Tania Leal who targets the construct of proficiency in SLA research. Building on insights gathered from her review of papers in *Studies in Second Language Acquisition*, Leal tackles two methodological issues: (1) the treatment of proficiency as a variable in SLA which she argues must be seen as a continuous construct rather than be grouped into arbitrary categories such as “beginner”, “intermediate”, “advanced,” and (2) the related question of sampling practices which should span as wide a proficiency spectrum as possible so as to avoid reduced statistical power and “spuriously significant results” (p. 74).

Part 2 investigates coding practices and concretely illustrates how changes in coding decisions impact the nature of the findings made about SLA development. This section includes two chapters.

- Chapter 4, by Megan Solon, revisits one of her prior studies (Solon, 2017) where she explored the acquisition of the lateral phoneme (i.e., /l/) in the spoken data of English-speaking learners of L2 Spanish. She had initially coded her data according to a pre-existing binary coding scheme (called ‘F2 values’), but subsequently recoded the same learner sample using a bottom-up approach: rather than relying on pre-determined labels, she created her own descriptive categories of lateral production based on what could objectively be observed in the learners’ speech. This method enabled several hitherto hidden developmental patterns to be uncovered, and supports the methodological claim that by studying learner data on the basis of a narrow selection of ‘a priori’ measures, one may overlook other emerging developmental patterns which are not necessarily captured by pre-selected measures. In other words, Solon urges SLA researchers to let the data speak “to hear what they might be telling us on their own” (p. 89).
- In Chapter 5, Amanda Edmonds and Aarnes Gudmestad reanalyzed two of their previous studies about the development of future time in L2 French (Edmonds & Gudmestad, 2015; Edmonds, Gudmestad & Donaldson, 2017). They recoded previously collected learner data samples according to two variables, namely temporal distance (‘near’ or ‘distant’ future), and adverbial specification (the presence or absence of an explicit time adverbial). The aim was to determine whether the level of detail with which these variables are coded influences the developmental findings concerning use of the future in L2 French. They showed that finer-grained analyses revealed more nuanced findings about the investigated features, thereby further emphasizing the methodological point made in Chapter 4 about the importance of data coding decisions and their impact on emerging L2 developmental profiles. Crucially, the authors insist on the need for much more explicitness in the description of L2 data coding decisions (in the form of publicly-shared descriptive guides, for instance) to enable studies to “engage in dialogue” (p. 127) with each other.

Part 3 describes specific approaches to SLA and is made up of three chapters about corpus linguistics, concept-oriented analysis and the use of naturalistic data, respectively.

- In Chapter 6, Nicole Tracy-Ventura and Amanda Huensch showcase the potential of using longitudinal learner corpora to capture L2 development and plead for such data to be made publicly available. The chapter details the impressive Language and Social Networks Abroad Project (LANGSNAP) where a group of university language majors (studying L2 French and Spanish) were followed over the course of several years. The resulting LANGSNAP corpus (oral and written data, c. 700 000 words) is used to study the effects of SLA variables (e.g., social networks, personality) on processes of attrition/maintenance/development in L2 acquisition. Given the efforts and budget necessary to the compilation of such databases, the authors urge for more longitudinal corpus data to be publicly shared and indicate how transcription and annotation can be made suitable for public sharing. The LANGSNAP data collected from 2011-2013 were generously donated to the CHILDES system (MacWhinney, 2000).
- Chapter 7 was written by Kathleen Bardovi-Harlig and is a critical account of concept-oriented analysis. The author illustrates how this method constitutes interlanguage analysis proper in that it

“treats the emergent learner variety as an independent system as envisioned by Selinker (1972)” (p. 188). Concept-oriented analysis relies on the prior identification of a concept (e.g. temporality, politeness, plurality, etc.), the detection and coding of the various learner linguistic means of expressing this concept (aka an “inventory”), how they change over time and what the relationship is between the different means of expression (“balance”). The constructs of “concept”, “inventory”, “change” and “balance” are thus central to this methodology. The concept-oriented approach is, in essence, a firmly linguistic analysis, one that is reminiscent of “in-depth case studies” (p. 189) rather than heavily dependent on “statistical inferences and comparison” (p. 189). The object of study is thus very much the “journey of development” (p. 184).

- The eighth and final chapter of the volume by Rachel Shively considers the challenges and opportunities of using naturalistic data in L2 pragmatics research. Shively shows how studying real social interactions with their “genuine affective, relational and material consequences” (p. 201) can broaden the scope of L2 pragmatic development and empirically reflect what learners actually do out ‘in the wild’. In her own database, the author focused on six United-States-based students who spent one semester in Spain and asked them to record a number of interactions (1) in their host families at dinner time, (2) with a Spanish-speaking age peer, and (3) in service encounters in local retail establishments. She also resorted to data triangulation and complemented the above data with language-learning journal entries, questionnaires about the learners’ experiences in Spain and one-on-one interviews with the learners themselves as well as their host families and peers. Revealingly, Shively shows that using naturalistic data enables, even forces, L2 pragmatics research to go beyond the study of the traditional speech acts (e.g., requests, refusals, invitations, etc.). In fact, her database contained very few instances of these speech acts in real-life social interactions, but rather revealed the presence of much more widely used features which have not been as actively investigated to-date, e.g., listener responses, speaker and listener assessments, teasing, and L2 humor development (p. 213).

Given the compellingly exemplary nature of each of the contributions in this volume, the short summaries provided above admittedly only do them partly justice. In the remainder of this review, I wish to put forward a (non-exhaustive) number of key methodological issues which the volume collectively points to and which I believe constitute some of the main contributions it makes to reflections about methodological practices in SLA research.

The first methodological take-away concerns construct definition and operationalization (e.g., Li, Leal). Chapter 3 by Leal about L2 proficiency deserves to be singled out as a truly thought-provoking contribution in this respect. The author rightly claims that L2 proficiency is such an “important piece of the puzzle in the study of L2 acquisition” (p. 70) that efforts must be made to improve its measurement across L2 studies. Not only does she explicitly deplore the heterogeneous ways in which L2 proficiency has been operationalized in SLA (e.g., institutional status, in-house or standardized tests, self-ratings, cloze tests, expert rating of individual productions, etc.) (see also Thomas 2006 for a critical review of proficiency

operationalization methods), she firmly condemns the practice which involves dichotomizing proficiency into discrete categories such as “beginner”, “intermediate” or “advanced”. Admittedly, the publication of the *Common European Framework of Reference for Languages* (CEFR) (Council of Europe, 2001) with its seemingly neat set of six distinct proficiency levels (A1, A2, B1, B2, C1, C2) has certainly contributed to reinforcing this methodological practice, as testified by the many research papers and language teaching manuals that rely on this system. Tellingly, Leal calls such dichotomous categories “artificial bins” (p. 75) in the sense that they are convenient and appealing in their simplicity but do not reflect the theoretical reality which is that proficiency is, in its essence, an intrinsically continuous and dynamic concept that develops slowly and progressively over time. This chapter literally shakes up our understanding of L2 proficiency and reminds us of what it truly is, namely a continuous concept that should not be put into ‘artificial bins’ simply for convenience’s sake. There is hope for improved methodological practice in this domain, however: innovative tools such as Comproved (<https://comproved.com/en/>), for instance, enable researchers to treat proficiency as a graded variable. Such tools rely on the practice of comparative judgement which involves comparing two learner texts and deciding which one is ‘better’. By comparing all texts multiple times, a reliable ranking is created. Learner texts are therefore not labelled as ‘intermediate’ or ‘advanced’, but are rather situated on an overall ranking, thus reflecting the continuous nature of proficiency assessments. In this vein, I wish to also refer the reader to the Crowdsourcing Language Assessment Project (CLAP) (Paquot, Rubin & Vandeweerd, 2022) carried out at UCLouvain (Belgium), which specifically relies on crowdsourcing tools similar to Comproved to study how they can contribute to improving L2 proficiency assessment practices. I strongly urge all researchers who are directly or indirectly concerned with proficiency-related methodological issues to take Leal’s arguments on board and start departing from dichotomization practices.

Second, praise is due to the SLA research field for opening up to the opportunities afforded by learner corpus research (LCR), and for the impressive efforts that have been put into compiling longitudinal naturalistic databases of L2 learner data to answer specific questions about SLA development (Tracy-Ventura & Huensch; Shively). The volume details the development of the LANGSNAP longitudinal learner corpus for learners of L2 French and Spanish (Tracy-Ventura & Huensch) as well as the naturalistic database compiled by Shively to study L2 Spanish development in learners abroad. I was truly impressed by the sheer amount of time, dedication, and ongoing critical reflection necessary to the compilation and maintenance of such databases— as well as the invaluable return on investment for longitudinal research agendas. To give but one example of the concrete commitment that goes into such work, the researchers travelled abroad several times



to personally maintain a rapport with the participants during their stay and limit data attrition. Shively additionally complemented her data with other sources of information such as journal entries or personal interviews, thereby illustrating the complementary insights that can be gained from combining naturalistic and other data types in a “methodological pluralism” approach (Gilquin, 2021, p. 135). This being said, I would also have appreciated a chapter dedicated to the existing longitudinal *and* cross-sectional corpora actively compiled by the LCR community¹ to illustrate how they have participated in SLA research agendas to-date. This would have further proved the point made in the volume that “once two quite separate fields, SLA and LCR have recently begun to come together in the joint effort to understand L2 development” (Tracy-Ventura & Huench, p. 153). I am thinking, for instance, of the *Longitudinal Database of Learner English* (LONGDALE) (Meunier, 2016), the *International Corpus of Learner English* (ICLE) (Granger *et al.*, 2020), or the *EF Cambridge Open Language Database* (EFCAMDAT) (<https://philarion.mml.cam.ac.uk/>) (Alexopoulou *et al.*, 2016), which have used LCR methods and tools to investigate SLA research questions. The new and publicly shared *Corpus of English as a Foreign Language* (COREFL) (<http://corefl.learnercorpora.com/>) (Lozano, Díaz-Negrillo & Callies, 2020) can now also be added to this list. At the moment, one may have the impression that SLA researchers develop their own learner corpora which fit their specific research purposes without yet making optimal use of the databases expertly developed by the LCR community. The reader may also be interested in consulting the recently published Routledge *Handbook of Second Language Acquisition and Corpora* (edited by Tracy-Ventura & Paquot, 2021) which takes stock of how learner corpora have been used in SLA research with the aim of encouraging theoretical and methodological cross-fertilisation between the two fields.

A third methodological take-away to stand out from this volume concerns the self-replication studies carried out by Solon and Edmonds & Gudmestad who revisited and recoded their own data samples and had the intellectual humility and honesty to show that there are several valid ways of approaching the same dataset with a tangible impact on L2 developmental results. Such self-replication work is very rare (dare I say almost non-existent) in SLA and, as such, constitutes a notable contribution of this volume. I hope that the compelling messages provided in these studies, i.e., that L2 data coding choices matter, that they crucially need to be made explicit in written guides, and that one should let the data speak, will encourage further initiatives of this kind. This issue is linked to the fourth major point to stand out from this volume, namely

¹ See the Learner Corpus Bibliography compiled by the Learner Corpus Association for a detailed overview of existing learner corpora around the world: <https://uclouvain.be/en/research-institutes/ilc/cecl/learner-corpora-around-the-world.html>



the call to improve methodological practices by participating in the open science movement. In the conclusion, Emma Marsden and Luke Plonsky, two forethinking figures in SLA methodological advancement and its relationship with open science, rightly stress that the SLA community has been rather slow in recognising the benefits of open science. This, they explain, partly justifies the lack of replication studies in SLA: without a transparent and publicly-shared outline of the research design and methods used in individual studies, replicating these objectively becomes an impossible challenge. The open science movement has slowly but encouragingly been gaining ground in the last few years, however. Marsden and Plonsky list representative initiatives which encourage open sharing in SLA, such as (1) the IRIS database (<https://www.iris-database.org/>) (Marsden & Mackey, 2014) which consists of freely downloadable peer-reviewed SLA materials, or (2) the open-source R statistical software. Since the publication of the current volume, other relevant open-source resources have been made available to the SLA community, such as for instance the TBLT Language Learning Task Bank (<https://tblt.indiana.edu/>) (Gurzynski-Weiss, n.d.), a central location for sharing language learning tasks or the Tools for Second Language Speech Research and Teaching ([Home - SLA Speech Tools \(sla-speech-tools.com\)](https://www.sla-speech-tools.com/)) (Mora-Plaza *et al.*, 2022). The message about the need to engage in the open-source movement also needs to be extended to the LCR community. While a small number of learner corpora have been made publicly available (e.g., the LANGSNAP corpus in this volume as well as the the EFCAMDAT and COREFL corpora), this still very much constitutes the exception rather the rule. Poor learner corpus availability hinders the active use of the same corpus resources by a wide variety of research teams, all of whom could collectively and distinctly add to findings about learner language.

This volume includes numerous other strong points such as the robust literature reviews carried out by Taguchi and Li who transparently outline the selection criteria used for the papers included in their databases. The chapter by Bardovi-Harlig about concept-oriented analysis is a highly appealing and convincing contribution, one which firmly analyses the interlanguage in its own right. It is my belief that this approach deserves more visibility in SLA research work. Other strong methodological points are peppered throughout the volume, such as the enlightening discussions on the ethics of naturalistic data collection (Shively), the (in)appropriacy of comparisons with target language norms (Bardovi-Harlig), or the statistical treatment of L2 proficiency (Leal), to name but a few. On a constructive note, perhaps the two chapters about interlanguage pragmatics (Chapters 1 & 8) could have been merged to make room for a contribution on a different aspect of SLA or, as suggested above, a chapter about the methodological opportunities and challenges offered by the learner corpora developed by the LCR community.



This volume is a most instructive and thought-provoking academic piece of work, with methods as its central focus. Each chapter truly leaves a mark on the reader, forcing active critical (re)thinking about our own methodological practices. I hope the book reaches a wide readership. It certainly deserves it.

Referencias

- Alexopoulou, Theodora; Michel, Marije; Murakami, Akira; Meurers, Detmar (2017). Task effects on linguistic complexity and accuracy: A large-scale learner corpus analysis employing natural language processing techniques. *Language Learning*, 67(S1), 180-208. <https://doi.org/10.1111/lang.12232>
- Byrnes, Heidi. (2013). Notes from the editor. *The Modern Language Journal*, 97, 825-827.
- Council of Europe. (2001). *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*. Council of Europe. <https://www.coe.int/en/web/common-european-framework-reference-languages>
- Edmonds, Amanda; Gudmestad, Aarnes (2015). What the present can tell us about the future: A variationist analysis of future-time expression in native and nonnative French. *Language, Interaction, Acquisition*, 6(1). 15-41. <https://doi.org/10.1075/lia.6.1.01edm>
- Edmonds, Amanda; Gudmestad, Aarnes; Donaldson, Bryan (2017). A concept-oriented analysis of future-time reference in native and near-native Hexagonal French. *Journal of French Language Studies*, 27(3), 381-404. <https://doi.org/10.1017/S0959269516000259>
- Gilquin, Gaëtanelle (2021). Combining Learner Corpora and Experimental Methods. In Nicole Tracy-Ventura, & Magali Paquot (Eds.), *Handbook of SLA and Corpora* (pp.133-144). Routledge.
- Granger, Sylviane; Dupont, Maïté; Meunier, Fanny; Naets, Hubert; Paquot, Magali (2020). *The International Corpus of Learner English* (Version 3). Presses universitaires de Louvain. <http://hdl.handle.net/2078.1/229877>
- Gurzynski-Weiss, Laura; IATBLT (n.d.). The TBLT Language Learning Task Bank. <https://tblt.indiana.edu>.
- Lozano, Cristóbal; Díaz-Negrillo, Ana; Callies, Marcus (2020). Designing and compiling a learner corpus of written and spoken narratives: COREFL. In Christiane Bongartz & Jacopo Torregrossa (Eds.), *What's in a Narrative? Variation in Story-Telling at the Interface between Language and Literacy* (pp. 21-46). Peter Lang. <https://doi.org/10.3726/978-3-653-05182-7>
- Marsden, Emma; Mackey, Alison (2014). IRIS: a new resource for second language research. *Linguistic Approaches to Bilingualism*, 4(1), 125-130. <https://doi.org/10.1075/lab.4.1.05mar>
- MacWhinney, Brian (2000). *The CHILDES project: Tools for analyzing talk: Transcription format and programs* (3rd ed.). Lawrence Erlbaum Associates Publishers.
- Meunier, Fanny (2016). Introduction to the LONGDALE Project. In Erik Castello; Katherine Ackerley; Francesca Coccetta (Eds.), *Studies in Learner Corpus Linguistics. Research and Applications for Foreign Language Teaching and Assessment* (pp. 123-126). Peter Lang. <https://doi.org/10.3726/978-3-0351-0736-4>
- Mora-Plaza, Ingrid; Saito, Kazuya; Suzukida, Yui; Dewaele, Jean-Marc; Tierney, Adam (2022). *Tools for second language speech research and teaching*. <http://sla-speech-tools.com/>
- Paquot, Magali; Rubin, Rachel; Vandeweerd, Nathan (2022). Crowdsourced Adaptive Comparative Judgment: A Community-Based Solution for Proficiency Rating. *Language Learning*, 72(3), 853-885. <https://doi.org/10.1111/lang.12498>



- Solon, Megan (2017). Do learners lighten up? Phonetic and allophonic acquisition of Spanish /l/ by English-speaking learners. *Studies in Second Language Acquisition*, 39, 801-832. <https://www.jstor.org/stable/26331052>
- Thomas, Margaret (2006). Research synthesis and historiography: The case of assessment in Second Language Proficiency. In John M. Norris & Lourdes Ortega (Eds.), *Synthesizing Research on Language Learning and Teaching* (pp. 279-298). John Benjamins. <https://doi.org/10.1075/llt.13.13tho>
- Tracy-Ventura, Nicole; Paquot, Magali (Eds.) (2021). *The Routledge Handbook of Second Language Acquisition and Corpora*. Routledge. <https://doi.org/10.4324/9781351137904>

