Learner Corpora in Language Testing and Assessment, Marcus Callies and Sandra Götz (eds). John Benjamins Publishing Company, Amsterdam (2015). 220 pp., ISBN: 978 90 272 0378 6.

Karim Sadeghi [a, *], Teymour Rahmati Kelahsarayi [a]

[a] Urmia University, Iran

Learner Corpora in Language Testing and Assessment is an edited volume in John Benjamins' *Studies in Corpus Linguistics* series. In the introduction to the volume, the editors specify their main objective as identifying the value of learner corpora in bringing transparency to the assessment of second/foreign language (L2) proficiency. It is also suggested that the operationalization of L2 proficiency using such internationally recognized frameworks as the Common European Framework of Reference for Languages (CEFR) be complemented by language specific lexical and grammatical descriptors derived from the analysis of learner corpus data. The volume has been organized into two sections each containing four articles. The contributions in section one present new corpus resources, tools, and methods specifically developed for language testing and assessment (LTA) purposes. The articles in section two deal with the problem of defining proficiency based on such vague criteria as institutional status and attempt to propose data-driven approaches to the assessment of proficiency.

In the first article, Rolf Kreyer points out that the reason learner corpora have failed to play a due part in LTA is rooted in their raw data, which are of cross-sectional nature representing advanced learners, and their flat structure annotation model, in which texts and annotations are stored in the same file thwarting the possibility of presenting more than one correction. He, then, introduces Marburg Corpus of Intermediate Learner English (MILE), which includes longitudinal data on intermediate learners. Regarding annotation, MILE uses both a text internal system and a text external multi-layered stand-off model in which annotations are stored independently from the texts to which they refer. MILE contributes to a data-driven approach to L2 proficiency

* Corresponding author: Urmia University, Iran
Email address: k.sadeghi@urmia.ac.ir

conceptualization by providing more text types at different levels of formality compared to only argumentative texts offered by most other corpora.

The second article by Gamallo, Garcia, del Rio, and Gonzalez focuses on an automatic error detection and correction tool called 'Avalingua'. The software has been developed to place learners in A1-C2 CEFR proficiency levels. The input texts are first scrutinized by a lexical module and then fed into spelling and syntactic modules. The spelling module automatically corrects words identified by the lexical module while the syntactic module detects and corrects erroneous syntactic relations. Finally, the student model scores the text based on the number of errors and text length. The modular structure of Avalingua allows it to be used for any language provided that the appropriate database is fed into the software. Avalingua enjoys an F-score (proportion between precision and recall) of 76% (compared to the lower F-score of 42% by other similar software).

In their contribution to the volume, Nordrum and Erikson deal with the challenge of writing and understanding data commentaries on visual materials by students of different disciplines and how this is addressed in a small specialized corpus which specifically focuses on 198 data commentaries extracted from 10 master's theses and 13 research papers published in three peer-reviewed international journals in the field of applied chemistry. The data commentaries are annotated for both move structures and lexico-grammatical features with the aim of developing "students' ability to self-assess their use of data commentaries" (p. 70) through different types of teacher-designed activities.

The last contribution in section one introduces a computer-mediated communication (CMC) corpus which contains 2023 comments on news stories posted online in a teacher created website. For the purposes of comparability, the corpus also includes a collection of over 23000 native speaker comments on news events in the BBC website 'Have Your Say'. A major goal of the designers is to assign proficiency levels to learner texts using binary 'yes/no' decisions by a group of raters for the concepts of complexity, accuracy, and fluency. The preliminary ratings reveal that most comments fall at levels 2 (31%), 1 (16%), and 3 (15%).

The first article in section two introduces English Vocabulary Profile (EVP) in which CEFR A1-C2 proficiency levels are assigned not to texts but to different definitions of individual words. In her article, Agniezka Leńok-Szymańska describes a study on the association between EVP-based and human raters' assigned CEFR levels to learners' texts in 90 A1-B2 level essays from the International Corpus of Crosslinguistic Interlanguage. The results reveal that EVP descriptions of text levels can be taken to be reliable. The study is among the first attempts to complement CEFR functionally characterized proficiency levels with text/data-driven descriptors.

In one out of the only two contributions devoted to speaking, Pascual pérez-Paredes and María Sánchez-Tornel report a study in which the linguistic features of the language produced by Spanish learners of English on a picture description component of the Louvain International Database of Spoken English Interlanguage (LINDSEI) were compared to the language produced by native speakers (NSs) on a similar task in interviews drawn from the Louvain Corpus of Native English Conversation. The data in the two corpora were first part-of-speech tagged and then explored using

multidimensional analysis, which is sensitive to co-occurring features. The findings reveal that the linguistic features used most frequently by NSs are less frequent in learners' speech, and vice-versa.

The article by María Belén Díez-Bedmar describes her study on the use of English articles for which a learner corpus of 26 essays written by Spanish learners of English was compiled. The essays were assigned to CEFR A2-B2 proficiency levels by two experienced raters and the criterial features of articles use at each level were investigated. Generally, the findings show that the criterial features of B2 level learners' productions are correct and effective use of English articles, in particular the zero article. Lower level learners' language, however, is characterized by the overuse of the definite article leading to the ineffective selection of the zero article. The research results bear witness to the potential benefits of corpus-based text/data-driven approach to LTA.

In the last contribution to the volume, Sandra Götz presents a study on verb-tense errors in the speech of German advanced learners of English, the data for which were drawn from the German sub-section of LINDSEI. To this end, 50 interviews were error-tagged by three experts with native speaker norms as point of reference. The findings reveal that the error category with the highest frequency is grammatical and related to verb tenses. The learners were, then, divided into three proficiency groups based on the number of errors. These results show that as proficiency increases, the categorical types and frequency of errors in learners' speech decrease. It is in this final article that samples of some CEFR functional characterizations of proficiency levels complemented by corpus-based data-driven lexico-grammatical descriptors are presented.

The volume can be of great interest to a wide range of readership among experts in corpus and computational linguistics, language assessment, L2 teaching, and ESP/EAP. Nevertheless, the potential benefits of the corpora and methods introduced in the volume have not yet been empirically investigated for real LTA purposes. Moreover, some of the methods (e.g., proficiency level assignment to CMC learner corpora texts) are based on such already challenged concepts as native speaker norms (Firth & Wagner, 2007; Holliday, 2008). The methods, which are also meant to be simple enough to allow self-assessment, are too subjective and complex to be implemented by learners. Furthermore, there are some practical considerations (e.g., hiring four raters as in the case of EVP project) which cannot be afforded by most LTA institutions.

Regarding content, it is recommended that the future editions of the volume include some more articles on the application of learner corpora in assessing speaking so as to achieve its aim of highlighting "the benefits and potential of using learner corpora for the testing and assessment of L2 proficiency in both speaking and writing" (p. 1).

The volume, however, enjoys a well-written introduction and a flexible organization in that readers can enter the book anywhere matching their own area of interest. Moreover, most of the contributions are examples of both new tools and data-driven approaches to proficiency assignment and readers with a clear understanding of the general aims of the volume can focus on these two themes in any article.

To sum up, the principal ideas of the volume are the necessity of developing tailor-made learner corpora for the purposes of LTA (Park, 2014) and the embodiment of lexico-grammatical features

in L2 proficiency conceptualization. *Learner Corpora in Language Testing and Assessment* may be regarded as a giant and fast forward step toward enhancing the role of learner corpora in LTA and opens new windows for rating writing and speaking in proficiency tests.

References

Firth, A., & Wagner, J. (2007). Second/foreign language learning as a social accomplishment: Elaborations on a receonceptualized SLA. *The Modern Language Journal, 91*, 800-819.

Holliday, A. (2008). Standards of English and politics of inclusion. *Language Teaching, 41*, 119-130.

Park, K. (2014). Corpora and Language Assessment: The State of the Art. *Language Assessment Quarterly, 11*(1), 27-44.