LESSIA M. KOTSYUK

National University of Ostroh Academy, Ostroh, Ukraine
lesya.kotsyuk@oa.edu.ua

ENGLISH LANGUAGE ERROR ANALYSIS OF THE WRITTEN TEXTS PRODUCED BY UKRAINIAN LEARNERS: DATA COLLECTION

Abstract

Recently, the studies of second language acquisition have tended to focus on learners errors as they help to predict the difficulties involved in acquiring a second language. Thus, teachers can be made aware of the difficult areas to be encountered by the students and pay special attention and devote emphasis to them. The research goals of the article are to define what error analysis is and how it is important in L2 teaching process, to state the significance of corpus studies in identifying of different types of errors and mistakes, to provide the results of error analysis of the corpus of written texts produced by Ukrainian learners. In this article, major types of errors in English as a second language for Ukrainian students are mentioned.

Keywords: corpus; data; English language; error; error analysis; Ukrainian learner

Introduction

Being a young independent state, Ukraine looks for the effective ways of integration into the European community. English became an essential part of the educational process in every university. Ostroh Academy National University chose English as a second working language along with the official Ukrainian language. The students use English during the classes, communicate with foreign professors and guests, write scientific papers, apply for programs abroad. The research group of lexicographic laboratory LEXILAB initiated creating of a Learner English Corpus to become a part of ICLE project and to represent samples of the Ukrainian English. While the data is being collected and organized, the research group of professors and students already make use of the Corpus to inform curriculum design at the College of Romance and Germanic Languages.

Error analysis in the process of L2 teaching

Errors are not anymore taken as bad signs of nonperformance, in teaching or learning, that should be eliminated at any cost; rather, they are considered to become a necessary part of language learning process. More than that there exists and functions a special type of linguistic analysis that focuses on the errors learners make. David Crystal (Crystal, 2003, p. 165) defines Error Analysis (EA) as a “technique for identifying, classifying and systematically interpreting the unacceptable forms produced by someone learning a foreign language, using any of the principles and procedures provided by linguistics”. Furthermore, M. Keshavarz (Keshavarz, 2012, p. 168) claims that EA is “a procedure used by both researchers and teachers which involves collecting samples of learner language, identifying errors, classifying them according to their nature and causes, and evaluating their seriousness”.

An analysis of errors in L2 learning surrounding is significant in three different ways (Selinker, 1969; Corder, 1981).

- (1) *To the teacher.* If a teacher initiates a complex permanent analysis, of whether the learners progress in their studies and, therefore, what remains for them to learn, it can acknowledge both the advances and defeats of the teaching programme. If a teacher understands what is hard to learn, and where the most of mistakes come from, he can add more efforts in the classroom practicing it with his students. According to S. K. Sharma (Sharma, 1980) “Error analysis can thus provide a strong support to remedial teaching”.
- (2) *To the researcher.* EA provides the researcher with information about the process of language learning and acquisition, the methods or strategies the student is using on his way to learn a language.
- (3) *To the learner.* Making of errors serves a device the learner uses in order to learn. If students understand what is hard to learn, they will pay attention to the grammar morphemes, the little words, and endings of words, not only when studying, but also when speaking.

In general, Error analysis has several implications for the handling of learner’s errors in the classroom: (1) Devising remedial measures; (2) Preparing a sequence of target language items in classrooms and textbooks with the difficult items coming after the easier ones; (3) Making suggestions about the nature or strategies of second language learning employed by both first and second language learners.

Models for error analysis

The establisher of EA Corder S. P. (Corder, 1973, 1975, 1981) has worked out the procedure for error analysis. Other scientists elaborated on this model: R. Ellis (Ellis, 1994), S. Gass and L. Selinker (Gass & Selinker, 2008).

Thus, the model for Error Analysis includes three stages: (1) Data collection — the selection of a corpus of language, written or oral one, the identification of errors; errors classification and quantification; (2) Description — a grammatical analysis of each error and the sources, and (3) Explanation (the ultimate object of error analysis) — explanation of different types of errors.

The significance of corpus studies in language error analysis

The tradition of language error analysis of the texts produced by learners goes back to 1960–70s. Some not very big collections of texts, non-electronic, variegated were

looked through. Only certain types of mistakes were fixed and extracted, the rest of the material was not taken into account. Modern learner corpora in contrast have a whole set of essential advantages: the amount of data is much bigger, electronic form, accurate criterion of data collection, etc. They provide the possibility not only to fix language errors but to make conclusions as to learners' speech in general and look critically at existing teaching methods, syllabuses, and teaching materials. Some works prove that corpora may provide data, which may further alter what is taught (D. Mindt (Mindt, 1996), S. Hunston (Hunston, 2002), T. McEnery and R. Xiao (McEnery & Xiao, 2011)).

The research group of lexicographic laboratory LEXILAB initiated creating of Ostroh Academy Corpus of Learner English (OACLE) to become a part of ICLE project and to represent samples of the Ukrainian English. The process of corpus creation is divided into several phases. The first step is to collect the essays of the students of the College of Romance and Germanic Languages. Every student can donate one or several essays, the number of words in each can range from 100 to 1000 words. The essays collected were of the some type — argumentative (a) written as a home task and allow using additional materials), or (b) written in class and does not allow using additional materials. At the moment 165 essays of (a)-type and 245 of (b)-type are taken from the students, whose native language is Ukrainian, they are at their first, second, third and the fourth year of studying at the College. All the samples were coded according to the system proposed by ICLE project.

The development of such a corpus appeared to be useful not only to participate in the international project but to inform the administration of the College as to typical mistakes, peculiarities of vocabulary choice, grammatical organization of the writings of the students.

Case study

As it was mentioned above, learner corpora can be used for the linguistic analysis, the aim of which is to identify the main lexical and syntactical errors of the students in the process of foreign language acquisition. It helps spot the frequency of different types of language errors, types of contexts in which a particular mistake was made, and develop effective plans and methods to improve foreign language teaching.

The research group of lexicographic laboratory LEXILAB set the aim to identify English language areas of difficulty for the students of the College of Romance and Germanic Languages in Ostroh Academy National University, which could help in (1) determining the sequence of presentation of target items in textbooks and classroom, with the difficult items following the easier ones; (2) deciding the relative degree of emphasis, explanation and practice required in putting across various items in the target language; (3) devising remedial lessons and exercises; and finally, (4) selecting items for testing the learner's proficiency.

Performing the experiment and collecting data is the starting point of the process of completing the experiment. Thus, according to the traditional framework of Error Analysis, three stages of the first phase were established.

(1) Collection of data

OACLE was the only source of data at this stage. The essays of students were selected. To be selected the essay should have met the following requirements: written by a student whose native language is Ukrainian; produced in class surrounding without using reference papers (dictionaries, textbooks, etc.) or web based error analysis tools (spellcheckers, grammar checkers, etc.). The task set for the students was the following:

“In the space below, write about yourself. Say where you come from, where you live, and give a little bit of information about your family and friends as well as your hobbies and interests. Finally, say where, when and how long you have been learning English. Also say how and why you would like to improve your English” (American inside out, n.d.).

The study dealt with written performance data from 245 students whose proficiency levels were previously assessed, varyingly, as level A2 (pre-intermediate), level B1 (intermediate), level B2 (upper-intermediate), level C1 (advanced), and level C2 (proficiency). Since the numbers of test takers in levels A2 and C1, C2 were low, they were not employed for the study. The total number of essays further analyzed was 217.

(2) Identification of errors

As far as OACLE is an untagged corpus the errors were detected with the help of Grammarly software. This grammar check service has a high Accuracy Score and recognizes a variety of common grammatical errors (Grammarly Review). Though Grammarly should not be considered the final word in spelling, syntax, and grammar we trusted it for the sake of this particular investigation. The errors found by the software serve the basis for the tag sets of corpus annotation.

(3) Classification into error types

Below are some examples of the grammatical errors identified in the sentences produced by the Ukrainian students of English. They can be classified into errors in spelling, grammar, punctuation, structure, and style.

Spelling Errors group falls into several subgroups, such as Contextual Spelling, Misspelled Words, Misspelled Grammatical Forms, Missing Hyphen in a Number, Lowercase Pronoun. The Contextual Spelling subgroup is represented mainly with mistakes due to inattentive printing of the text, i.e. *My mother works in a bank ans (and) is very intelligent and determind person. What abous (about) my hobbies? Familly*. The examples of typically misspelled words are: *quatation* (quotation), *valleyball* (volleyball), *sturted* (started), *lern* (learn), *conected* (connected), *proffesiional* (professional), *alredy* (already), *espesially* (especially), *hovewer* (however); misspelled grammatical forms: *studing* (studying), missing hyphen in a number: *twenty two* (twenty-two), *My siblings are second(-)year student right now*.

Grammar Errors group contains the subgroups of errors in article use, incorrect verb form, misuse of quantifiers, incorrect noun number, and misuse of modifiers.

The Article Use subgroup can be exemplified by cases of missing article, i.e. *I have (a) close friend. We have (the) tradition in our family, we have movinights and dinner together.*; missing article before proper names, i.e. *Now I am a student*

of (the) *National University of Ostrog Academy*.; using the definite article instead of indefinite, i.e. *I am the (a) third year student of the National university of Ostroh academy. I think the (a) good way to do this is to immence yourself in the English-speaking atmosphere.*; using the indefinite article with a plural noun, i.e. *Each day I take a pictures of a sky.* Some examples of incorrect verb form use are: *We like (to) spend weekends together, as we all study in different cities and do not have possibility to meet more often. To conclud, I want to say we have work(ed) hard to develop our inner world and knowledge of language, and we should do it all the time.*

The Misuse of Quantifiers subgroup in Grammar Errors group contains the cases where the students write singular quantifier with a plural noun, i.e. ... *and that is why I see my parents not each weekends.*, or singular noun after plural quantifier, for example, *My name is Tania and I was born at one of the most freeziest day in January, 1993 in little village.* Also, there are cases of using this with a plural noun: *This (These) both styles are very different but both worth admiring.*

Noun number can also be the source of errors in students' writings. There were noticed some sentences in which singular noun was used with the plural number: *I spend 3 month(s) in Canada.* Another error case in this subgroup is forming plural instead of possessive, i.e. *My parents(') names are Eygeniya and Vasil.*

Erroneous double comparative was placed in the subgroup of Misuse of Modifiers: *I have a dream to go to London and live there for some time to feel their life and to be more better in English.*

Punctuational errors formed a separate **group** in which such types of problems were noticed: 1) Comma misuse within a clause, missing comma after introductory phrase in particular: *In winter(,) we like spending long cold evenings by watching some amazing film or just reading a fascinating book. In my native town(,) I have a lot of friends.*; 2) Missing comma in a compound sentence, i.e. *My name is O.(,) and I come from Rivne region. Poland and Canada helped me to proved my English skills(,) but I know that it is not the end.*; 3) Unnecessary comma in a complex sentence: *Some of them I can see every day, because we are studing together, going out and living. We like spend weekends together, as we all study in different cities and do not have possibility to meet more often.*; 4) Comma splice: *My father is a mechanic by profession, (;) he likes reading books and watching television. I come from Sarny, (;) that is a Rivne region. ;* 5) Incorrect punctuation with quotation mark: *"Family is not an important thing, it's everything(.)"*.

Structure Errors group contains problems that deal with misplaced words or phrases. The examples are as follows: *With this town connected a lot of precious memories. For me is better to have a few friends, that I can trust all my worries and feelings.*

The last group of errors found in students' writings is **Style Errors**. These are the problems connected with, firstly, improper formatting, such as incorrect spacing with punctuation: *I am easy-going,friendly and kind person, so I have a lot of friends in Sarny and Ostrogh.*, or capitalization at the start of a sentence: *that is my cup of tea.* The second type of stylistic errors is wordiness, i.e. *In order to improve my foreign languages I will pay more attention to self-education, listening different records and watch films in English and German, attend different courses,*

because if I want to have a good workplace I have to be a proffesiional. / In general, my family is very small, It is me and my mother, but we have so tight relations with our relatives that I can even say I have maybe the biggest family in this town or even in whole district. As well, students erroneously use redundant words: Actually they all are my old school friends whom I keep in touch with. / I really hope so. / However, it was a hard time of building our home and all my family was doing their own work, especialy when there were a lot of work to do in the fild.

Conclusion

The purpose of the study was to give a look at the first steps in the process of error analysis of the corpus of Ukrainian students' essays written in English. The results of the present survey indicate that error analysis is a complicated process consisting of several stages. Data collection is the primary step in this process and presupposes the collection of essays, error identification and classification into its types. As the result, 217 essays were collected, grammar checked and errors were grouped into spelling, grammar, punctuation, structure, and style types. Each type in the paper was exemplified with sentences from students' works. The findings will be useful for the teachers of English in Ukraine in compiling teaching materials, students of English in the process of learning, researchers and developers of automated grammar checkers. Future research will deal with deep analysis of the sources of errors, their context interpretation and their dependence on learner proficiency level and native language.

References

- American inside out. (n.d.). Retrieved 25 November 2014, from <http://www.insideout.net/new/files/2010/10/full-placement-test.doc>
- Corder, S. P. (1973). *Introducing applied linguistics*. Harmondsworth: Penguin Books Ltd.
- Corder, S. P. (1975). Error analysis. In J. P. B. Allen & S. P. Corder (Eds.), *Papers in applied linguistics: The Edinburgh course in applied linguistics* (Vol. 3: *Techniques in applied linguistics*). Oxford: Oxford University Press.
- Corder, S. P. (1981). *The significance of learner's errors*. Oxford: Oxford University Press.
- Crystal, D. (2003). *A dictionary of linguistics and phonetics* (5th ed.). London: Blackwell. <http://doi.wiley.com/10.1002/9781444302776>
- Ellis, R. & Barkhuizen, G. (2005). *Analysing learner language*. Oxford: Oxford University Press.
- Ellis, R. (1994). *The study of second language acquisition*. Oxford: OUP.
- Gass, S. & Selinker, L. (2008). *Second language acquisition: An introductory course* (2nd ed.). New York: Routledge.
- Grammarly Review. (n.d.). Retrieved 25 November 2014, from <http://online-grammar-check-review.toptenreviews.com/grammarly-review.html>
- Hunston, S. (2002). *Corpora in applied linguistics*. Cambridge: Cambridge University Press. <http://ebooks.cambridge.org/ref/id/CB09781139524773>
- James, C. (2013). *Errors in language learning and use: Exploring error analysis*. London: Routledge.

- Keshavarz, M. (2012). *Contrastive analysis and error analysis* (2nd ed.). Tehran: Rahamana Press.
- McEney, T. & Xiao, R. (2011). What corpora can offer in language teaching and learning. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (Vol. 2). New York: Routledge.
- Mindt, D. (1996). English corpus linguistics and the foreign language teaching syllabus. In J. Thomas & M. Short (Eds.), *Using corpora for language research*. Harlow: Longman.
- Selinker, L. (1969). Interlanguage. *IRAL — International Review of Applied Linguistics in Language Teaching*, 10(1–4). <http://doi.org/10.1515/iral.1972.10.1-4.209>
- Sharma, S. K. (1980). Practical and theoretical consideration involved in error analysis. *Indian Journal of Applied Linguistics*, 6, 74–83.

Acknowledgment

This work was supported by a core funding for statutory activities from the Ministry of Education and Science of Ukraine.

The author declares that she has no competing interests.

This is an Open Access article distributed under the terms of the Creative Commons Attribution 3.0 PL License (<http://creativecommons.org/licenses/by/3.0/pl/>), which permits redistribution, commercial and non-commercial, provided that the article is properly cited.

© The Author 2015

Publisher: Institute of Slavic Studies, PAS, University of Silesia & The Slavic Foundation