

Errors in Learner Corpus: Pedagogical Implications for KFL Instruction

Author:

Shin, Seong-Chul

Publication details:

Working Paper No. Vol 4, No 10
KAREC Discussion Papers

Publication Date:

2003

DOI:

<https://doi.org/10.26190/unsworks/451>

License:

<https://creativecommons.org/licenses/by-nc-nd/3.0/au/>

Link to license to see what you are allowed to do with this resource.

Downloaded from <http://hdl.handle.net/1959.4/39405> in <https://unsworks.unsw.edu.au> on 2024-04-18

Errors in Learner Corpus: Pedagogical Implications for KFL Instruction

Seong-Chul Shin

(University of New South Wales)

1. Introduction

The analysis of learner errors has been one of the first methods used by researchers to investigate second-language acquisition (SLA) or learning. The error analysis (EA) research, which was initiated to supplant the weaknesses of contrastive analysis (CA), was considerably popular in the 1960s and 1970s. At these times, EA was closely associated with the work of Corder (e.g. 1967, 1971a, 1974), who made significant contributions to the development of EA methodology and whose work is still widely referred to. Though EA lost popularity as a result of the perceived weaknesses including the procedural problems and the limitations in the scope, there were continuing signs of SLA research through EA in the 1980s and onward (for example, Davies 1983; Taylor 1986; Lennon 1991). In fact, the starting point for EA in Korean was from the late 1980s (for example, Sohn H-M 1986; Lee S-J 1987), and a good number of EA case studies in Korean has been carried out over the last ten years or so, particularly from the mid 90s to the recent years (for example, Kim M-O 1994; Wang H-S 1995; Che O-Y 1997; Lee E-K 1999; Kim Y-M 2000; Shin S-C 2001, 2002; Lee J-H 2001, 2002; Kim C-S and Nam K-C 2002; Kim J-S 2002; Ko S-J 2002; Han S-H 2002).

This increasing popularity in recent years is not surprising in that it coincides with the increasing number of tertiary institutions offering Korean as a foreign language (KFL) programs in Korea and overseas in that period. It is now reported that a comprehensive file of errors called 'Yonsei Corpus of Learner Language' has been compiled, and some studies emphasize the importance of machine readable data (MRD), with suggestions for a computerised analysis system (e.g. You S-H 2001; Seo S-K et al 2002). In spite of such a considerable research output in EA research in Korean, there seems to be insufficient or inadequate information about pedagogical implications of error for teaching of KFL. This paper intends to address what implications the research findings have for teaching and make suggestions about how they can or should be utilised for

pedagogical purposes. Along with this aim, it will look at, in passing, some methodological problems that appear to be present in the current EA studies in Korean.

2. Significance of Errors and Pedagogical Implications

Corder (1967) states that a learner's errors could be significant in three ways: 1) they provide the teacher with information about how much is learnt; 2) they provide the researcher with evidence about how language is learnt; and 3) they are indispensable devices for the learner to learn the target language. The first function is a practical one concerning the need for remedial action (e.g. remedial teaching and remedial materials); the second function is a theoretical one concerning the methodology of investigating the second language learning process; and the third function is a combination of both, where a learner repeats constructing and re-constructing a set of hypotheses and then discovers the system of the target language.

All these roles are inter-related and have significant pedagogical implications for teaching. For pedagogical purposes, we first need to know what are the main areas of weakness in the learner's Korean. This can be achieved by both qualitative and quantitative approaches. To identify the principal learning difficulties we need a qualitative linguistic classification of errors and quantitative information of the relative frequency of each error type. Before this analysis process, however, we will need to reach an agreement about the definitions used in the classification. Unfortunately linguists of the Korean language disagree on the functional and grammatical categories and accordingly researchers have employed different terms in their description and classification of errors. Also, we will need a more adequate classification and categorisation system.

3. Qualitative Classification

A systematic technique to classify and categorise the errors will enable both researchers and practising instructors to identify and describe error types more adequately. Corder (1975, 1981) suggests a matrix for the error categorisation: 1) according to the phenomenon: omission, addition, selection/substitution and ordering; 2) at linguistic level: orthography or phonology (graphological/phonological), morphology or syntax

(grammatical), vocabulary (lexico-semantic); 3) according to systems: vowel or consonant systems, tense, aspect, mood and so on. Sohn H-M (1986) analyses his Korean data by using such linguistic levels and categories as: 1) orthography; 2) lexicon (nouns, verbs, adverbials, Sino-Korean collocations; 3) morphology (modifier endings, *l*-related, infinitive suffix, copula-related, irregular verbs, case particle allomorphs and others; 4) syntax (word order, case markers and delimiters, tense, negation, clause conjunction, nominalization, complementation); 5) sociolinguistics and pragmatics.

In the meantime, Lennon (1991) proposes two new dimensions of error: domain and extent. Error domain refers to the breadth of context (word, phrase, clause, sentence or discourse), which is adopted to determine whether error has occurred. Extent refers to the size of the linguistic unit (morpheme, word, phrase, clause or sentence) that needs deleting, replacing, re-ordering or supplying for repair of an erroneous production. Lee J-H (2002a) shows a comprehensive chart of error categorisation according to the cause and end results. For the end results three different ways are suggested: 1) by category (pronunciation, grammar, vocabulary and others; 2) by phenomenon (substitution, omission, addition); 3) by the erroneous degree (whole/global, partial/local).

Though individual researchers may adopt approaches that suit their research purpose and plan, it will be more efficient and systematic in research and teaching if we can obtain a consensus that can provide us with a simple but systematic and comprehensive classification tool. This paper does not intend to argue for a particular approach to error classification but for preliminary discussion, one might wish to propose a possible model for Korean.

4. Quantitative Statement of Frequency

Once classifications are satisfactory and descriptions are completed, it is useful or necessary for practising teachers to have a quantitative statement of the frequency of each type of error so that they can make themselves aware of the principal areas of difficulties that learners experience, thereby paying more attention to those areas or taking appropriate action for improvement. It will be useful for teachers if the frequency statement includes information such as: 1) frequency of errors at each linguistic level or category; 2) frequency of errors by pattern or phenomenon; 3) high frequency universal

errors irrespective of the learner's language background and learning stage; 4) high frequency universal errors according to the learner's language background; 5) high frequency universal errors according to the learner's learning stage; and further 6) the relative frequency of errors based on teaching method and materials that were employed. Although this sort of information is highly statistical, it will certainly give teachers a 'big picture' about absolute and comparative areas of difficulty and at what stage the learners are in their learning career. Further, it will provide teachers with useful information about where to emphasise and what kind of input (i.e. method and materials) is needed to address the problem area.

Previous studies have some common features in terms of high frequency at each linguistic category. For example, in Kim M-O (1994), where errors produced by English and Japanese speakers learning Korean were investigated, the most frequent in occurrence were case particle errors with 40.5% and 36.5%, respectively. This case-marking problem has been widely found in other studies with varying degrees of percentage (e.g. Sohn 1986; Lee E-K 1999; Kang H-J 2000; Kim J-S et al 2002; Kim M-O 2002; Lee J-H 2002b). Shin S-C (2003 forthcoming), where English native speakers studying Korean in Korea and Australia were compared, also confirms that the most frequent grammatical errors come from the erroneous use of case particles (43.6% and 52.7% respectively).

At other levels of description also, there seems to be quite common features in high frequency error type and pattern, for example, *ae* for *e* or *e* for *ae*, *eo* for *o* or *o* for *eo*, *j* for *ch* or *ch* for *j*, and *s* for *ss* or *ss* for *s* at orthographic or phonological level (e.g. Kim M-O 2001; Shin S-C 2001; Lee J-H 2002b); semantic similarity, overgeneralisation and literal translation as the cause of some main lexical errors (e.g. Wang H-S 1995; Shin S-C 2002); confusions between *-go* and *-aseo/-eoseo*, and *-aseo/-eoseo* and *(-eu)nikka* as key conjunctive errors (e.g. Lee S-J 1987; Kim J-S 2002; Lee J-H 2002b; Shin S-C 2003 forthcoming). In this way, one might feel that it is necessary to provide a comprehensive frequency percentage list of common learner errors. A study of such percentage values is important in that it provides instructors and researchers with an insight into the relative significance of a given error type in the total context. On the basis of the percentage

values of the errors, we may be able to decide an order of preference for different drills when we plan remedial work.

5. Understanding the Nature of Errors

Although statistical information gives us a good insight into the relative significance of an error, it is not sufficient for an effective remedial teaching. It is necessary to conduct a deeper analysis of the particular error. We need to understand the nature of the learner's difficulties and give an efficient explanation of the cause of the error. This is because we can only do something about the problem areas in a systematic way when we understand why the error has occurred. A number of studies have identified the linguistic and psycholinguistic sources of the errors. For example, Richard (1971b) identifies three sources or causes of errors: 1) interference; 2) intralingual (overgeneralization, ignorance of rule restrictions, incomplete application of rules, and falsely hypothesized concepts); 3) developmental errors. Similarly, Dulay and Burt (1974b) classified their error data into three categories of cause: 1) developmental; 2) interference (transfer); and 3) unique (e.g. induced). Sohn (1986) reports that interlingual transfer errors are more universal than intralingual errors in his data. Lee J-H (2002a) suggests more detailed categories of cause: 1) mother tongue interference (negative transfer); 2) target language interference (overgeneralization, incomplete application); and 3) pedagogical cause (teaching materials and teaching methods).

There seems to be a general agreement in the way of explaining the cause of the error. Understanding the cause or source of the error gives us an insight into the second language learning process, and this is directly or indirectly related to the provision of effective language teaching materials and methods. Given that the learner tests the hypothesised system of the target language on the basis of the language input or information given to him/her, it is critical for him/her to receive complete and clear information rather than incomplete and misleading information. In fact, it seems that a large portion of the learner errors in Korean are caused by ambiguous or simplistic information, and in this regard this paper intends to emphasise the relative significance of the induced or pedagogical cause of the error mentioned above. We need to give the

learners sufficient data or explanation so that they can formulate correct hypotheses, thereby forming a better or native-like 'approximate system'.

6. Remedial Measures: Materials

Once we know the nature of the error, we may be able to undertake appropriate remedial steps. Corder (1975) suggests pedagogical relevance of error analysis in three categories but it seems that the direct pedagogical implications are twofold: the design of remedial syllabuses or materials, and the design of pedagogical grammars. Given that a number of studies have found the similarities between L1 and L2 acquisition, particularly in syntactic structures, it will be possible or desirable to elaborate on the teaching order and the developmental sequence in the KFL syllabus and teaching materials by utilising the information obtained from the learner's approximate system.

Thus far, there is little or no experiment work which has actually tried out the acquisition order based on error analyses in Korean, though there are some suggestions or in-house experiments for effective teaching of particular language items or tasks (e.g. see Kang H-J 2000, Kim C-S and Nam K-C 2002). In designing materials on individual linguistic items, we will need to pay maximum attention to the main areas of difficulty in remedial drills and give a secondary importance to areas with less problems. Remedial drills can be more effective if spoken exercises or activities come first, followed by written practice of drills, which will help reinforce the familiar or mastered patterns. As remedial teaching requires a clinical approach, the main drill excises can be constructed in four ways: repetition, substitution, transformation and recombination. By doing so, the learner will not simply repeat and substitute the same or similar item, but also re-think or re-express his/her ideas in an appropriate way and manipulate one or more additional items with the already mastered items.

7. Remedial Measures: Grammar

The writing of pedagogical grammars must also reflect the language learning strategies and processes that have been discovered from the learner's interlanguage data. The main challenge here is threefold: 1) how we can give the learner simple but clear-cut and complete information or explanation; 2) how effectively we can present the linguistic

materials to the learner; and 3) how we will enable the learner to practise the grammatical elements effectively. As far as I am aware, there exist no pedagogic grammars based upon error analysis and experimental work in Korean.

Beyond drills and grammar, we will also need to consider the organizational perspective in both spoken and written language – whether the sentence or speech is adequately developed, whether a point is clearly constructed, whether the information is relevant to the topic and whether the paragraphs are coherently ordered.

8. Concluding Remarks

In conclusion, while we need to be cautious about so-called the learner-centred approach, we will need to look at the effectiveness and significance of the bottom-up approach, rather than relying on the top-to-down *Dr-knows-it-better* approach, as has been the case. Information from the learner's language has significant values not only in remedial teaching but also in ordinary teaching, and when we have more sophisticated learner data, the learner can be more sophisticated about their production.

Selected References:

- Corder, S.P.1967. 'The significance of learner's errors'. *International Reviews of Applied Linguistics* 5, and also in Corder.1981.
- Corder, S.P. 1971a. 'Idiosyncratic dialects and error analysis'. *International Review of Applied Linguistics* 9: 149-59. Also, in Corder 1981.
- Corder, S.P.1975. 'Error analysis, interlanguage and second language acquisition'. *Language Teaching and Linguistics Abstracts* Vol 8.
- Corder, S.P. 1981. *Error Analysis and Interlanguage*, Oxford: Oxford University Press.
- Davies, E. 1983. 'Error evaluation: the importance of viewpoint'. *English Language Teaching Journal* 37:304-11.
- Dulay, H. and M. Burt. 1974b. 'Errors and strategies in child second language acquisition'. *TESOL Quarterly* 8: 129-36.
- Ellis, R. 1994. *The Study of Second Language Acquisition*. Oxford: Oxford University Press.
- James, C.1998. *Errors in Language Learning and Use*. London: Longman.
- Geroe, H. 1972. *Common Errors in Language Learning: Insights from English*. Rowley, Mass.: Newbury House.
- Larsen-Freeman, D. and M. Long. 1991. *An Introduction to Second Language Acquisition Research*. London: Longman.
- Lennon, P. 1991. 'Error: Some problems of definition, identification, and distinction'. *Applied Linguistics* 12: 2.
- Nemser, W. 1971. 'Approximate systems of foreign language learners'. *IRAL*, Vol. 9 No. 2. Also in Richards (ed.) 1974.

- Norris, J. 1983. *Language Learners and their Errors*. London: Macmillan Publishers.
- Richards, J. C. 1971b. 'A non-contrastive approach to error analysis', in Richards R. C. (ed.). 1974. *Error Analysis: Perspectives on Second Language Acquisition*. London: Longman.
- Selinker, L. 1972. 'Interlanguage'. IRAL 10: 209-31. Also in Richards (ed.) 1974.
- Shin, S-C. 2001. 'An error analysis: Orthographic errors produced by Australian KFL learners'. *Teaching Korean as a foreign language*. Vol 25-26. Yonsei University.
- Shin, S-C. 2002. 'Australian students' lexical errors in Korean: Type, frequency and cause'. *Journal of Korean Language Education* Vol 13, No 1.
- Sohn, H.M. 1986. 'Patterns of American students' composition errors in Korean', *Linguistics Expeditions*. Seoul: Hanshin Publishing Co.
- Taylor, B. 1975. 'The use of overgeneralization and transfer learning strategies by elementary and intermediate students of ESL'. *Language Learning* 25: 73-107.
- 고석주. 2002. 학습자 말뭉치에서 조사 오류의 특징. 외국어로서의 한국어 교육. 27
- 김미옥. 1994. 한국어 학습에 나타난 오류 분석. 한국말 교육 5. 국제한국어교육학회.
- 김미옥. 2002. 학습 단계에 따른 한국어 학습자 오류의 통계적 분석. 외국어로서의 한국어 교육 27 집.
- 김유미. 2000. 학습자 말뭉치를 이용한 한국어 학습자 오류 분석 연구. 연세대학교 교육대학원 석사학위논문.
- 김정숙. 남기춘. 2002. 영어권 한국어 학습자의 조사 사용 오류 분석과 교육 방법. 한국어 교육 13-1.
- 김중섭. 2002. 한국어 학습자의 연결 어미 오류 양상에 관한 연구. 한국어 교육 13-2
- 서상규. 유현경. 남윤진. 2002. 한국어 학습자 말뭉치와 한국어 교육. 한국어 교육 13-1.
- 안주호. 2002. 한국어 교육에서의 [원인] 연결어미에 대하여. 한국어 교육 13-2.
- 왕혜숙. 1995. 영어화자의 한국어 작문에 나타난 어휘상 오류 분석. 이중언어학회지 12. 이중언어학회.
- 유석훈. 2001. 외국어로서의 한국어 학습자 말뭉치 구축의 필요성과 자료 분석. 한국어 교육 12-1.
- 이순자. 1987. 문법의 분석: 영어화자의 한국어 습득과정에 일어난 오류를 중심으로. 연세대학교 교육대학원 석사학위논문.
- 이은경. 1999. 한국어 학습자의 조사 사용에 나타난 오류 분석. 연세대학교 대학원 석사학위논문.
- 이정희. 2001. 한국어 학습자의 시제 오류 연구. 이중언어학 18. 이중언어학회.
- 이정희. 2002. 한국어 학습자의 표현 오류 연구. 경희대학교 박사학위논문.
- 최우영. 1997. 외국어로서의 한국어 학습자의 오류에 대한 연구. 이화여대 석사학위논문.
- 한송화. 2002. 한국어 학습자의 오류 분석 - 전성 어미에서의 오류. 외국어로서의 한국어 교육 27 집.

Korean Studies
University of New South Wales,
Sydney, 2052, Australia
e-mail: s.shin@unsw.edu.au
phone: (+612) 9385-3724