LEXICAL ERROR ANALYSIS IN THE WRITTEN PRODUCTION OF STUDENTS OF ENGLISH AS A SECOND LANGUAGE: A PILOT STUDY

Begoña Sarrionandia Gurtubay
Universidad del País Vasco/Euskal Herriko Unibertsitatea

This research tries to carry on, by using the model of error análisis (EA), a transversal study of the lexical errors committed by students of English as a Second Language in a written corpus at university level. The main aims are to offer an inventory of lexical errors made by Spanish speaking students of English. Secondly, this study aims at presenting a taxonomy of errors at the lexical level, based on linguistic and descriptive criteria, that is, dealing with one of the linguistic subsystems, the lexical one, and etiological criteria, that is, to describe the cause of error, intralinguistic errors and interlinguistic ones. Besides, this analysis categorizes intralinguistic and interlinguistic lexical errors according to the strategy used by the learner. Interlinguistic errors would be classified in transference of the signifier and transference of the meaning. In the same way, and depending on the strategy used, intralinguistic errors can be categorized into simplification, analogy and linguistic development. Finally, the results will be used to offer a statistical analysis which shows error frequency and, consequently, some conclusions will be extracted.

Key words: error analysis, contrastive analysis, lexical competence.

Esta investigación pretende llevar a cabo un estudio transversal de los errores léxicos realizados por los estudiantes universitarios de inglés como segunda lengua en un corpus escrito. Para ello, se utilizará el modelo de análisis de errores (EA). Los principales objetivos son ofrecer una relación de errores léxicos cometidos por los estudiantes de inglés hispanoparlantes. En segundo lugar, este trabajo tiene como finalidad presentar una taxonomía de los errores a nivel léxico, y los criterios etiológicos, es decir, describir las causas de los errores, tanto los intralingüísticos como los interlingüísticos. Según esto, este análisis clasifica los errores en errores léxicos intralingüísticos e

Submission: 16/02/2009 – Acceptance: 11/06/2009

interlingüísticos, dependiendo de la estrategia usada por los estudiantes. Los errores interlingüísticos pueden ser de dos tipos: la transferencia del significante y del significado. Del mismo modo, según la estrategia utilizada, los errores intralingüísticos se pueden catalogar como simplificación, analogía y desarrollo lingüístico. Por último, se usarán los resultados para ofrecer un análisis estadístico que muestre la frecuencia de error y que, por lo tanto, permita llegar a algunas conclusiones.

Palabras clave: análisis de errores, análisis contrastivo, competencia léxica.

ERROR ANALYSIS: STATUS QUESTIONS

First of all, and in order to understand the motivation on which this research is based, I shall place the method of Error Analysis (EA) in the frame of Applied Linguistics. This method of analysis has its precedents in the so-called Contrastive Analysis (CA), dated in the 40s and 50s (Friers 1945). By using this behaviourist analysis, L2 learners errors can be avoided by comparing the two linguistic systems inside the process in a systematic way, native language or L1 and second language or L2. The starting hypothesis states that if both systems present similarities at any level (phonetic, phonological, grammatical or semantic) the acquisition of the target language would be improved and the possibility of making errors would decrease, and, on the opposite, if the linguistic systems immersed in the process present a greater number of differences the acquisition of the L2 would be more difficult and the probability of making errors would be greater, or in other words, error frequency would increase.

With regard to this, Charles Friers states:

> The most efficient materials are those that are based upon a scientific description of the language to be learned, carefully compared with a parallel description of the native language of the learner. (Friers 1945:9)

In the same sense, Lado claims the following:

> These elements that are similar to his native language will be simple for him, and those elements that are different will be difficult. (Lado 1957:29)

To sum up, some authors defend the so-called Contrastive Analysis Hypothesis (CAH), which states that whenever two languages present similarities there is a positive transference, and while both languages are different there is a negative transference or interference.

As a consequence of the former statements, the Contrastive Analysis Model would be summarized as follows:

1. Structural description of L1 and L2.
2. Comparison of descriptions.
3. List of equivalent structures.
4. Difficulty hierarchy.
5. Prediction and description of learning difficulties.
6. Didactic material.

So CA tries to deal with the problems of second language learning, but it does not describe how students learn a second language. Due to this point of view, one of the consequences of this analysis model is the negative attitude towards errors, considered as a failure of communication. A second pedagogical implication is the adoption of the audio-lingual or grammar/translation method, that is, to avoid the interference of mother tongue. Errors are treated as inadequate answer to stimuli, which have to be corrected.

On the other hand, this theory about the learning process of languages does not avoid the fact that speakers of different languages make the same errors.

In the 60s, and after the revolutionary conceptions about language acquisition promulgated by Noam Chomsky who stated that the learner created grammatical rules based on the input, and the statement that claimed that learners of languages have an innate talent for acquiring languages (language acquisition device), authors like Ritchie (1967) rejected the method of CA, stating that speakers could take different learning or communicative strategies (competence or performance), because of the language acquisition device.

(1) Knowledge of L1 — knowledge of L2
L1 deviation — L2 deviation

(2) Transference Grammar
L1 → L2
L2 → L1

This means that the content of the course should not be based on the relationship between the native and the second language (transfer grammar or contrastive analysis), but on the relationship between the learners' innate linguistic knowledge. Thus, languages only differ in the deep structure, not in the surface one.

As a consequence, a new method of analysis of the process of linguistic acquisition was applied, the so called Interlanguage Analysis (Selinker 1972). This method tried to analyze the learners' idiosyncratic competence at a determinate level in a systematic way. This transitory competence is an

intermediate level of learning, which can be used to measure the learner's competence or describe the linguistic system the learner uses, so the learning process is seen as a creative and cognitive process.

In the 70s, Corder (1967) proposes that errors are evidence of learners' strategies, not inhibition signals, and defends the systematic study of errors in order to describe the syllabus built by the learners.

Consequently, EA proposes empirical and scientific procedures to analyze learner's production in a L2. This model of analysis stated that errors are not arbitrary, but, on the contrary, errors are a source of information of the systematic way in which a language is acquired. The procedure consists of six steps:

(i) Corpus selection of homogeneous groups,

(ii) Error identification,

(iii) Classification,

(iv) Explanation,

(v) Evaluation.

Corder distinguishes between errors and mistakes, or, in other words, systematic errors of competence and systematic errors of performance. Therefore, the great contribution of EA to the field of applied linguistics is the new conception about error, its decriminalization, and the increase of its status.

AIMS OF THE STUDY

This research tries to carry on, by using the model of EA, a transversal study of the lexical errors made by students of English as a Second Language in a written corpus at university level. Thus, the main aims are the following:

1. To offer an inventory of lexical errors made by Spanish students of English.
2. To present a taxonomy of errors at the lexical level, based on linguistic and descriptive criteria, that is, dealing with one of the linguistic subsystems (lexical), and etiological criteria, that is, to describe the cause or source of errors, intralinguistic errors and interlinguistic ones.
    i. Interlinguistic: Transference of the L1 sign (signifier or meaning) as it belongs to the L2.
    ii. Intralinguistic: Transference of the L2 sign (form) to other L2 sign, which is correct in isolation, but does not work in a particular context.

3. To categorize the intralinguistic and interlinguistic lexical errors according to the strategy chosen by the learner. Interlinguistic errors would be classified in the following way:

   i. Transference of the signifier: Transformation of a L1 signifier into a L2 form in a wrong manner. This phenomenon is also called lexical borrowing. This type of errors is divided into lexical or phonetic errors.

   ii. Transference of the meaning: Transformation of a L1 meaning into a similar L2 form. These types of errors are also called "false friends" or semantic borrowings.

   In the same way, and depending on the strategy used, intralinguistic errors are categorized into:

   i. Simplification: Forms that are part of the signifier. Erroneous deviations will be included.

   ii. Analogy: Forms that share part of the meaning. In the same way, inadequate semantic forms will be included (those that belong to the same semantic field).

   iii. Linguistic development: Application of false suppositions because of lack of knowledge of the adequate L2 form. Learner can apply substitution rules in an indirect way (periphrasis) or direct one (synonyms).

4. To offer a statistical analysis which shows error frequency and extracts conclusions from the statistical data.

5. To extract didactic implications after the analysis of the results. Concerning this, a hierarchy of error gravity according to their communicative effect will be offered. This hierarchy presents a system of bands and scales that could be a helpful tool to assess learners´ lexical competence.

CORPUS

Pilot groups that have provided the data for are composed by 67 university students from several degrees that study English as a Second Language at the Faculty of Letters (University of the Basque Country). These students have followed a formal instruction throughout one semester. Data for the corpus have been taken from the written examination. One of the sections of the examination is the writing of an essay (free topic). Learners are instructed to complete the

essay in 30 minutes, and the average extension is of 125 words. The examination was taken in November (2004) and June (2005).

METHOD OF ANALYSIS

The typology of errors consists of lexical errors (morphological and orthographic), dealing with the signifier, and semantic, related to the meaning. Discursive, structural, and syntactic errors have been collected, but not considered in the scope of this research, and will be included in following researches concerning this field. The total number of errors (378) has been categorized by applying the following taxonomy:

Interlinguistic
- Transference of signifier
  - Lexical level (lexical borrowing)
  - Phonetic / orthographic level (transference of phonetic/ orthographic rule)
- Transference of meaning
  - Semantic borrowing (heterosemantic terms)

Intralinguistic
- Simplification (erroneous derivations/lexical omissions)
- Analogy (inadequate semantic/orthographical forms)
- Linguistic Development (lack of knowledge/substitution)

CATEGORIZATION

Interlinguistic errors

(i) Transference of meaning:

i. Syntactic borrowing:

(1) × "Then he listened the story"

✓ "Then he listened to the story"

The learner makes a wrong selection because of negative influence: "escuchar algo" (× "to listen something") does not subcategorize a preposition in Spanish (subject [-human]), as it does in English.

ii. Semantic borrowing

(2) × "She tries to talk with her""

✓ "She tries to talk to her"

Direct translation because of negative transference of native language. In Spanish "hablar con" (× "talk with"), and in English "talk to". The occurrence (2) has been categorized as a semantic borrowing.

(3) × "The people was"
✓ "People were"

Transference of meaning in the verbal form. In occurrence (3) Spanish, L1 of the pilot groups, transfers the agreement of number [-plural], and, thus, the occurrence is wrong in English. Apart from this, the inclusion of the definite article is not adequate in English.

(ii)Transference of signifier

i. Lexical borrowing

(4) × "Pieces of paper"
✓ "Sheets of paper"

Negative influence of the Spanish "pedazos" ("pieces"). Lexical borrowing.

(5) × "The story appears unreal"
✓ "The story seems unreal"

The learner applies the similarity between Spanish form "parecer" ("seem") to English "appear".

(6) × "The conductor"
✓ "The driver"

Negative transference of the Spanish signifier "conductor".

(7) × "To pass our weekend"
✓ "To spend our weekend"

The Spanish signifier "pasar" [+time] is transferred in a negative way.

ii. Phonetic borrowing

Consonantal germination

(8) × "babysiter"
✓ "babysitter"

(9) × "ofered"
✓ "offered"

In Spanish orthography there in no germination.

Spellimg

(10) × "agly"
✓ "ugly"

(11) × "extrange"
✓ "strange"

(12) × "dificult"
✓ "difficult"

(13) × "caracter"
✓ "character"

Occurrences (10), (11), (12) and (13) show the transference of Spanish phonetic system to the English spelling of former signifiers.

Intralinguistic errors

(i) Simplification

(14) × "It was late, she dead"
✓ "It was late, she was dead"

Omission of lexical unity. Default application of passive voice. Misunderstanding or lack of knowledge of the adjective "dead" versus the verb "died".

(15) × "All people"
✓ "All the people"

Omission of lexical unity. "People" as subject without definite article in English. So, the learner makes an intralinguistic error when using the pronoun "all" as subject. Hyper generalization of the Spanish use of "all".

(16) × "Without lucky"
✓ "Without luck"

Erroneous semantic derivation. Use of the noun "luck" instead of using the adjective "lucky". The same phenomenon is applied to the occurrence (17):

(17) × "He tells him a really story"
✓ "He tells him a real story"

(18) × "He said his master could help him"
✓ "He asked his master to help him"

Misunderstanding of forms in L2, "say" instead of "ask".

(i) Analogy

(19) × "Every people"
✓ "All the people"

Analogous forms "every/all" in Spanish. Excessive application of the form "every". Lexical forms belonging to the same semantic field.

(20) × "she did not meet the house"
✓ "She did not find the house"

Lexical forms "meet/find" belonging to the same semantic field.

Occurrences (21), (22), and (23) show the same misunderstanding of analogous pairs, in this case time and local adverbs which share semantic features, as in "ago/before", "on/in", and "at/ ø":

(21) × "The girl has broken her neck years ago"
✓ "The girl has broken her neck years before"

(22) × "On the middle of the road"
✓ "In the middle of the road"

(23) × "At the third time"
✓ "The third time"

(24) × "He asked his master could help him"
✓ "He asked for his master to help him"

"Ask/ask for" analogous spelling forms. Omission of the preposition "for".

(25) × "To back at home"
✓ "To go back home"

Omission of the verb "to go". Excessive application of the preposition "at" in inadequate context.

(ii) Linguistic Development

(26) × "The return path"
✓ "The way back"

Erroneous synonym because of lack of knowledge. Direct substitution strategy.

(27) × "On winter"
✓ "In winter"

False supposition. Application of the preposition "on" when dealing with seasons instead of dates.

(28) × "At the end"
✓ "In the end"

(29) × "To bring into the world"
✓ "To give birth"

Periphrasis because of lack of knowledge.

ANALYSIS OF THE RESULTS

A total of 378 errors have been categorized. After analyzing the strategy used, 213 interlinguistic errors and 165 intralinguistic errors have been typed. So, by using the taxonomy and the categorization of the total number of errors the following conclusions have been extracted:

1. Learners of English as a Second Language with intermediate level of competence after a period of instruction present interlinguistic errors by transference in a greater frequency than intralinguistic errors. This conclusion is related to the statement that claims that L2 learners based their performance on their mother tongues.

Taking into account that this research is a cross linguistic study, it will measure the interlanguage of the pilot groups in a concrete phase of the learning process. A longitudinal analysis would show the trend of other similar researches (Linde 1992) that claim that learners will make a greater frequency of errors at intermediate levels of the learning process. This statement does not deny our results: 56.3% of interlinguistic errors and a percentage of 43.6% in intralinguistic. These results can be contrasted with similar studies (Dulay & Burt 1974), in which intralinguistic errors represent a small percentage of the sample.

In our case, we can conclude that there is a balance between the errors made due to the transference of the mother tongue, and those made because of the influence of the L2. A similar research (Sarrionandia 2006) concerning Spanish

speaking learners of Basque as L2, stated that the optimum level of intermediate competence presents a balance between the use of communicative strategies based on L1 (source of interlinguistic errors), and strategies based on L2 (source of intralinguistic errors).

2. Among the interlinguistic errors, the strategy of transference of meaning presents a greater percentage (56%, 121 errors), than errors made because the effect of transference of signifier (43%, 93 occurrences). Interlinguistic errors of the form or sign (signifier and meaning) are produced more at the semantic level than at the lexical one. Thus, evidence relates to the fact that L2 learner acquires a language at the semantic level at the first stages of the process of learning, and when this learner transfers the meaning to the L2, although the student does not know the adequate form, she decides to translate the meaning literally, at a sintagmatic level, or in isolation, at word level, The learner translates the meaning to a similar form of the L2 which is being acquired. The resulting form is called *false friend* or heterosemantic term.

3. When dealing with intralinguistic errors the percentage is not comparatively significant among the strategies used. It can be advanced that learners would make a greater percentage of errors categorized as errors of Linguistic Development, taking into account the level of interlanguage in which they are. However, this phenomenon is not produced, and learners use strategies of simplification and analogy at the same extent.

Consequently, we can deduce that the learner is aware of the distance that is between both linguistic systems (Spanish and English) inmersed in the process, and tends to use a strategy of avoidance when he/she does not know the term. In this way, learners achieve a positive transference with omission of the lexical element of the L2. Besides, learner also produces a strategy of ultracorrection, that is, learners think that the forms of both systems are too similar, and do not produce the adequate L2 form.

4. After recounting the total number of errors, we conclude that the areas of lexical difficulty that present a greater number of occurrences are the verb paradigms and prepositions (locative, directional, and time prepositions). Thus, learners have made a total of 65 errors in the field of verb paradigms and 27 errors in the choice of prepositions.

GRAVITY OF ERRORS

In Corder´s words (1974), gravity of error consists in the *mismatch* that is between the knowledge of a language and the needs of the linguistic situation.

In order to determine the grade of irritability, gravity, or *mismatch* of an error, this error has to be tested to a gradation of communicative impact. In order to carry out this assessment, and categorize learner's lexical competence, the assessor can use a system of bands and scales as a secondary criterion of assessment. Several authors (Linde 1992, Sarrionandia 2006) have stated that errors that avoid communication are those at the lexical level, just because they do perform this type of errors not only in isolation, but also at the syntagmatic and discursive levels. Thus, we propose a table of hierarchy of error gravity:

Type	**Orthographical**	**Lexical**
1st grade errors	Not serious deviation. Correct orthography. Reader does not present problems to recognize the form.	Not serious deviation. Reader does not present problems in substituting the form.
2nd grade errors	Not serious deviation. The form can be interpreted by the context.	Interpretable with the context.
3rd grade errors	Serious deviation. Difficult interpretation, even with the help of the context.	Difficult interpretation by the context.
4th grade errors	Serious deviation.	Important gravity. Reader presents serious problems in identifying the item.
5th grade errors	Impossible interpretation. Reader cannot interpret the meaning of the form.	Impossible to identify the form, even with the help of the context.

Table 1. Error gravity

DIDACTIC IMPLICATIONS

In this research we have followed the method of explaining the errors after they have been made, that is, the so called Weak Hypothesis of the Contrastive Analysis. This research has pretended to be a methodological proposal that allows to categorize the lexical competence of a pilot group of Spanish speaking students of English as a Second Language, under the perspective of Error Analysis.

However, we need to quote the possible didactic implications that can be derived from this study, and although each instructor can use this proposal to analyze materials, level of acquisition, and methods of instruction, we suggest the following chart in order to be taken as a an error therapy:

Identification: × To pass our weekend

Occurrence (token) (Type of error	**Reconstruction**	**Description**	**Classification**	**Source of error**
10 Transference of signifier	To spend our weekend	Literal translation. Spanish "pasar"	Lexical borrowing	Interlinguistic error

Table 2. Error treatment

The instructor can apply the chart (Table 2) to each of the errors that have been identified. Then, he/she can select the categories that have presented a greater number of occurrences, in order to avoid the collective error and determine which lexical category shows a greater level of difficulty for the learner. In this way, the instructor will take the adequate method in order to avoid or diminish their use by groups of similar characteristics.

Concerning the method quoted above, there are different procedures to deal with errors (Hendrickson 1984), that can be used by the instructor to carry out after the indexation of errors:

(i) To offer the correct form. This is a feedback corrective treatment.

(ii) Indirect treatment. The instructor evidences the error and does not offer solutions. Learners carry out a self-correction individually or in pairs or groups.

Authors like Cassany (1996) have supported this kind of procedures. The learner would know his/her deficiencies when making hypothesis about the rules of the interlanguage, and will achieve the adequate ways to solve them; and the instructor would control the effectiveness of his/her teaching methods, and the contents that have not been acquired by his/her learners.

Apart from the didactic implications of the corrective methods, the instructor can define the syllabus for similar groups of learners. The application of this method makes the instructor know the real learning situation, and he/she can immediately solve its supposed problems, such as the excess of contents or the existence of unsuccessful contents in the syllabus, which usually is an important source of errors, or, on the opposite, a default of necessary lexical contents for communicative effects.

Finally, we have to add that learners errors in L2 are indicators of learning problems in a foreign language, in our case, learning of English language, and this contribution has tried to diminish the lack of successful communicative effects at lexical level.

REFERENCES

Cassany, Daniel. *Reparar la escritura. Didáctica de la corrección de lo escrito*. Barcelona: Editorial Grao, 1996.

Corder, Pit. "The significance of language errors." *International Review of Applied Linguistics* 5 (1967): 161-171.

Corder, Pit. "Error Analysis", *The Edimburgh Course in Applied Linguistics*. Eds. John P. Allen and Steven Pit Corder, *Techniques in Applied Linguistics*. vol. 3. London: Oxford University Press, 1971.

Dulay, Heidi, & Marina Burt. "Natural sequences in child second language acquisition." *Language Learning* 24 (1974): 37-53.

Friers, Carol. *Teaching and Learning English as a Foreign Language*. Ann Arbor: University of Michigan Press, 1945.

Hendrickson, James. "The treatment of error in written work." *Composing in Second Language*. Ed. Sandra McKay. Rowley, Mass.: Newsbury House Publishers, 1984: 216-21

Linde, Angeles. 1992. "Errores en el uso de las formas verbales inglesas en la expresión escrita por alumnos universitarios." *Análisis de errores en inglés: tres casos prácticos*. Ed. Antonio Bueno. Granada: Universidad de Granada, 1992: 107-165.

Myles, Johanne. "Second language writing and research: the writing process and error analysis in student text." *Teaching English as a Second or Foreign Language. TESL-FL*. 6.2 (2002).

Santos Gargallo, Isabel. "El análisis de los errores en la interlengua del hablante no nativo." *Vademécum para la formación de profesores*. Eds. Jesús Sánchez Lobato & Isabel Santos Gargallo. Madrid: SGEL, 2004: 391-410.

Sarrionandia, Begoña. "Tipología de errores: análisis de un corpus oral y escrito entre hablantes euskaldunizados." *Ohienart. Cuadernos de Lengua y Literatura* 21 (2006): 467-479.

Selinker, Larry. "Interlanguage." *International Review of Applied Linguistics* 10 (1972): 219-231.