Analysis of graphic errors in the written productions of students of Spanish as a foreign language

Mónica Belda-Torrijos
Universidad Cardenal Herrera-CEU, CEU Universities, Spain

María Gloria García-Blay
Universidad Cardenal Herrera-CEU, CEU Universities, Spain

Linda Palfreeman
Universidad Cardenal Herrera-CEU, CEU Universities, Spain

Abstract

The fundamental focus of this research is the observation and analysis of the most frequent graphic errors in the learning of Spanish as a foreign language by students in a formal context, the classroom, and from a corpus of written language. The first global analysis of the study allowed us to observe which errors are the most common, as well as their frequency. The graphic errors are those that refer to punctuation and other signs, tildes, separation and union of words, alteration of the order of the letters, confusion of phonemes, omission of letters, addition of letters, confusion of graphemes for the same phoneme, graphemes and non-Castilian, capital letters. The descriptive study carried out will allow us to present first a quantitative analysis (with figures and percentages) of all the graphic errors observed, as well as some representative qualitative examples. This analysis has allowed us to ascertain which values are the most affected. We have carried out an error analysis based on grammatical category, this being a type of taxonomy that clearly demonstrates the linguistic competence of the student.

Keywords: graphic errors; punctuation; accent markers; phoneme confusion; confusion of graphemes; non-Spanish digraphs; capital letters.

Análisis de los errores gráficos en las producciones escritas de los estudiantes de español como lengua extranjera

Resumen

El enfoque fundamental de esta investigación es la observación y el análisis de los errores gráficos más frecuentes en el aprendizaje del español como lengua extranjera por parte de los estudiantes en un contexto formal, el aula, y a partir de un corpus de lengua escrita. El primer análisis global del estudio ha permitido observar qué errores son los más comunes, así como su frecuencia. Los errores gráficos son los que se refieren a la puntuación y otros signos, las tildes, la separación y unión de palabras, la alteración del orden de las letras, la confusión de fonemas, la omisión de letras, la adición de letras, la confusión de grafemas para un mismo fonema, los grafemas y las mayúsculas. El estudio descriptivo realizado nos permitirá presentar primero un análisis cuantitativo (con cifras y porcentajes) de todos los errores gráficos observados, así como algunos ejemplos cualitativos representativos. Este análisis nos ha permitido determinar cuáles son los valores más afectados. Hemos realizado un análisis de los errores en función de la categoría gramatical, siendo esta un tipo de taxonomía que demuestra claramente la competencia lingüística del alumno.

Palabras clave: errores gráficos; puntuación; marcadores de acentuación; confusión de fonemas; confusión de grafemas; dígrafos no españoles; mayúsculas.

Introduction

The acquisition/learning of an L2 has been one of the topics that has aroused a great deal of interest and has encouraged the development of various theories. Klein (1986) presents a synthesis of the main research approaches and orders them under four aspects: the identity hypothesis, the monitoring theory, the contrastive analysis model and the error analysis model.

The error analysis (EA) model aims to point out the areas of difficulty in learning a second language for a certain group of students with the same native language, therefore, it intends to establish an inventory of the most frequent errors and to assess their importance and severity. The use of a taxonomy assumes that a particular error has a specific source and that the specification of the source is a descriptive task, but one must be aware that the description of an error is somewhat different from the task of inferring the cause of said error.

There are different criteria depending on whether the objective of the analysis is grammatical competence or communicative competence. Below we show the different types of errors according to the following criteria:

Descriptive criteria

Depending on the way in which the surface structure of the instances is altered. Typology:
- Omission. This is a process that consists in suppressing morphemes and words that are not redundant. The omission of morphemes is caused by an incomplete knowledge of the rules of grammar, while the omission that affects lexical morphemes is usually due to lack of vocabulary.
- Addition. The result of adding morphemes or words that are redundant. The fundamental cause is an excessively faithful use of the rules due to linguistic hypercorrection.

- Incorrect training. It is the use of a word that has been formed or derived erroneously by adding or omitting of any of its morphemes. The cause is the ignorance or insecurity of the student who, in the absence of a synonym within their competence, tries to guess the form, making use of false analogies.
- Absence of sentence order. It is the incorrect placement of a morpheme or set of morphemes in the utterance, affecting the sentence order of the language system being learned.

Pedagogical Criterion

This criterion is based on the Chomskian distinction between competence and performance that led Corder (1974) to differentiate between: transient errors and systematic errors. The former are the result of circumstances at certain stages of the learning process and are not systematic. What causes them is usually a lack of adequacy in the teaching techniques or in the order of presentation of structures. The latter reflect a poor knowledge of the rules and characterize the language of a group of students. They are errors identifiable by contrastive studies between the native language and the target language.

Etiological-linguistic criteria

This classification is based on the concept of linguistic transfer. We can distinguish between interlinguistic errors and intralinguistic errors. Interlinguistic errors occur due to interference with the student's native language or with other languages that they have previously learned. Ringbon (1986) states that depending on the distance between L1 and L2, this will be the frequency index of the interference. Intralinguistic errors are those that refer to the effect of the target language itself on the learning process.

Grammatical Criteria

This type of taxonomy allows for the assessment of the grammatical competence of the student. The criteria are classified as:
- Phonological errors
- Spelling errors
- Morphological errors
- Syntactic errors
- Lexical errors
- Semantic errors
- Pragmatic errors

Communicative criteria

The classification of errors is based on those that hinder the transmission of the message. To establish the typology, criteria such as: acceptability, adequacy, intelligibility and irritation caused in the listener, etc. are used. Corder established two types of error: local and global. Local errors are those that affect individual elements of the sentence and have no impact on the breakdown of communication and global errors are those that affect the whole sentence and that due to their high level of unacceptability and impropriety cause a communication breakdown.

Methodology

Taking into account the methodological principles of corpus linguistics, we applied a working method that consisted of compiling a written corpus of texts written by different learners. The corpus on which the research is based is made up of a total of 766 texts written by university students of different levels. Each of the students wrote a personal diary and a critical commentary on their own individual blogs, as part of four classes per week. Then, the mistakes made by each student were recorded.

An error analysis was carried out based on grammatical category and adding discursive and graphic errors. This classification is based on the categories established Fernández (1997) and Santos (1993) in order to collect all errors and to create a broad classification with its corresponding labels. The classification will be part of the methodology, as well as being used for the analysis of the results. It is a type of taxonomy that clearly demonstrates the linguistic competence of the student. This custom designed tool was designed to establish a work method that would allow us to manage the information obtained and to classify it. We categorize and label errors as follows:

• ERRORES GRÁFICOS .... EGR
  Graphical errors

  » PUNTUACIÓN Y OTROS SIGNOS.................................... EGR_P
  Punctuation and other marks. Incorrect use of punctuation marks
    - Coma..EGR_P_com
     Comma
    - Punto y coma...................................EGR_P_puntcom
     Semi-colon
    - Punto ....EGR_P_punt
     Full stop
    - Guión ....EGR_P_gui
     Hyphen
    - Comillas...EGR_P_comill
     Quotation marks
    - Signos de interrogación....................EGR_P_interrog
     Interrogation marks. Placement of question marks only at the end of the sentence.
     The rules of writing in English cause confusion for the learner when writing in Spanish.
    - Signos de exclamación ....EGR_P_eclam
     Exclamation marks. Placing of exclamation marks only at the end of the sentence.
    - El guion y la división silábica.............EGR_P_guids
     The hyphen and syllabic division
    - Los dos puntos................................EGR_dpunt
     Colon
    - Los puntos suspensivos ....EGR_P_puntsp
     Ellipsis marks

  » LAS TILDES... EGR_tilde
   Accents. Errors associated with the incorrect use of the orthographic accent, or by its omission.
  » SEPARACIÓN Y UNIÓN DE PALABRAS....................................EGR_sup
   Separation and union of words. Error in the separation of words

» ALTERACIÓN DEL ORDEN DE LAS LETRAS...........................EGR_aol

  Alteration in the ordering of letters.

» CONFUSIÓN DE FONEMAS...EGR_cf

  Confusion of phonemes. Errors committed on substituting one letter for another that does not figure in the Word.

» OMISIÓN DE LETRAS ....EGR_ol

  Omission of letters. Error consisting in the omission of one or more letters in a word.

» ADICIÓN DE LETRAS...EGR_al

  Addition of letters. The erroneous adding of letters to a word.

» CONFUSIÓN DE GRAFEMAS PARA EL MISMO FONEMA.......EGR_cg

  Confusion of graphemes for the same phoneme. The student uses an incorrect grapheme for the phoneme referred to.

» GRAFEMAS Y DIGRADOS NO CASTELLANOS........................EGR_gdnc

  Non-castilian graphemes and digraphs. The use of graphemes and digraphs from another language.

» MAYÚSCULAS ....EGR_may

  Capital letters.

Results and discussion

In this article we treat the mistakes made by: punctuation and other marks, accents, separation and union of words, alteration of the order of letters, confusion of phonemes, omission of letters, addition of letters, confusion of graphemes for the same phoneme, non-Castilian graphemes and digraphs and incorrect use of capital letters. The 9,895 spelling and phonetic errors recorded represent 55.50% of the total errors collected in the compositions – an important number from the quantitative point of view.

Table 1. Graphical errors

GRAPHICAL ERRORS		TOTAL	PORCENTAJE
PUNCTUATION AND OTHER MARKS	Comma	526	5.32%
	Semi-colon	18	0.18%
	Full stop	0	0.00%
	Hyphen	1	0.01%
	Quotation marks	23	0.23%
	Interrogation mark	76	0.77%
	Exclamation mark	344	3.48%
	The script and syllabic division	0	0.00%
	Colon	2	0.02%
	Ellipsis points	1509	15.25%
Accents		4835	48.86%
Word separation and union		149	1.51%
Alteration in order of letters		258	2.61%
Confusion of phonemes		245	2.48%
Omission of letters		244	2.47%
Addition of letters		270	2.73%
Confusion of graphemes for the same phoneme		530	5.36%
Graphemes and non-Castilian digraphs		17	0.13%
Capital letters		848	8.57%

Authors' own elaboration

Within the section punctuation and other marks we observe the following error percentages:

Table 2. Figures and percentages of graphic errors

GRAPHICAL ERRORS			
		TOTAL	PORCENTAJE
PUNCTUATION AND OTHER MARKS	Comma	526	21.05%
	Semi-colon	18	0.72%
	Full stop	0	0.00%
	Hyphen	1	0.04%
	Quotation marks	23	0.92%
	Interrogation mark	76	3.04%
	Exclamation mark	344	13.77%
	The script and syllabic division	0	0.00%
	Colon	2	0.08%
	Ellipsis points	1509	60.38%

Authors' own elaboration

Of the 2,499 errors found at this point, the incorrect use of ellipsis, the comma and the exclamation marks, in that order, stand out. This high number of errors is explained by the lack of knowledge on the part of the learner of punctuation rules.

ACCENTS

Errors related to the incorrect use of the accent, either by omission, wrong choice or addition add up to 4,835, that is, 48.86% of the total spelling errors. Most of the errors have wrong uses where the tonic syllable is confused.
The following are some examples:
   • Omission: *obligo; *esta (verbo estar); *después; *ultima
   • Incorrect use: *hablarè; *frió

What clearly causes numerous errors is the ignorance on the part of the learners regarding the rule that the accent on monosyllables is a diacritical accent, and is used to distinguish words that have the same form but different meaning or different grammatical function.
*te (tea); *se (the verb 'to know'); *mas (adverb of quantity, 'more'); *mi (personal pronoun, 'my')

SEPARATION AND UNION OF WORDS

With respect to the separation and union of words, 149 errors were found. There is a loss of syllabic integrity due to ignorance:
   *playas bonitas y anfi teatros y edificios históricos; *Creo que a bajo no hay espacio para todos;
   *setenta porciento del total de la UE
   On the other hand, we observe numerous errors due to the separation of the conjunction 'because' that introduces cause or motive:
   *por que corría hacia la condenada; *el mundo entero espera mucho de él por que ha prometido cambios
   Incorrect separation of numbers:
   *Un viaje inolvidable: veinte dos personas, tres coches (dos coches de nueve placas; *Iraq en un plazo de diez y seis meses

Contraction of the definite article:

*Entonces hemos ido <u>a el</u> aeropuerto y hemos esperado; *Entonces hemos vuelto <u>a el</u> hotel a las once; *Hemos llegado <u>a el</u> aeropuerto a las nueve

CHANGING THE ORDER OF LETTERS

This element comprises a total of 258 errors, the most common errors being the separation of personal pronouns from direct and indirect objects and reflexive pronouns.

*quiero <u>le dar</u> la dirección de mi blog; *creído que hoy no puedo <u>me levantar</u>

PHONEME CONFUSION

Regarding the confusion of phonemes, 245 errors are found regarding the hesitation in the use of the phonemes e / i, o / u, e / a, o / a, e / o and vice versa:

*kilómetros y tengo que aprovechar al <u>fundo</u>; *He vuelto a casa a las 9. Hemos <u>cucinado</u> y comido; *entonces podemos <u>discubrir</u> muchas otras culturas

Some of the errors are caused by carelessness when representing the phonemes graphically, and are explained by the interference of the mother tongue or ignorance of the use of the appropriate phoneme.

OMISSION OF LETTERS

*<u>Necesto</u> el español para trabajar; *Por ejemplo las dos <u>gueras</u> sin sentido; *comer <u>churos</u> con chocolate caliente

ADDITION OF LETTERS

*diferente; *classe; * inteligente; *interessante

CONFUSION OF GRAPHEMES FOR THE SAME PHONEME

There are 530 errors concerning the confusion of graphemes for the same phoneme.

*es el <u>centre</u> de industrial lugar que textil <u>empezé;</u> lleno y todo el mundo <u>paresia</u> que se sentía bien; *A las <u>quatro</u> yo y İpek fuimos a Alzomora

GRAPHEMES AND NON-CASTILIAN DIGRAPHS

* que un <u>dá¬a empezo</u> y no va a terminar; *Tiene la nariz <u>pequeña</u>

CAPITAL LETTERS

The number of graphical errors with respect to capital letters is significant: 848 errors. We find errors in the first word of a text and in that which goes after a full stop, regardless of whether it is preceded by the opening of parentheses, quotation marks, interrogation or exclamation mark:

*<u>muchas</u> gracias a todos los personas; *<u>hoy</u> es muy importante para mi…

There are errors in the use of the lower case in proper names:

*Ella escucha música todos los días. Le gusta <u>julio iglesas</u>; *Se llaman <u>esin y sena. esin</u> tiene 25 años

The specific names of streets, urban spaces, roads and highways requires a capital letter as well as the name of commercial establishments. Errors such as the following have been found:

*en la plaza de la<u> reina</u> en una terraza soleado; *Esta noche voy a ir a <u>western</u> ;*Ahora voy al <u>mercadona</u>

The use of lowercase for proper names of countries, continents, cities and localities is repeated in numerous compositions:

*Por ejemplo, los <u>estados unidos</u> tienen una grandísima crisis financiera: mucho desempleo, recesión; *Hoy estoy muy feliz porque mañana voy a volver a <u>francia</u>

Unnecessary use of capital letters:

*Cuando <u>Llegamos </u>a Alicante, hicimos una vuelta; *casi veinte estudiantes erasmus (<u>Belgas</u>, turcos, polacos, lituanos,..

Conclusions

Ut The search for the causes that originate each error is extensive, but it is sometimes necessary to address the error in a more precise way and to facilitate its correction. In orthographic errors we observe mechanisms of neutralization of oppositions, hypercorrections, ignorance of the rules, interference of the L1 in graphemes and non-Castilian digraphs.

It is important to know how to distinguish between systematic and non-systematic errors. The errors that are due to the lapse of memory, physical disposition and psychological states are considered non-systematic and are of no particular interest with regard to the learning process of a language, be it L1 or L2, since they are normally correctable by the speakers themselves.

On the other hand, those errors that are the result of the underlying knowledge of the target language are systematic errors and show the language procedure that the learners are using. This systematic error/non-systematic error duality is an externalization of Chomskian discrimination between competence and performance, according to which systematic error is a deficiency of linguistic competence, while non-systematic error belongs to the level of action and occurs because of its improper use of linguistic knowledge, without indicating ignorance of the rules. It is to the systematic errors that the model in question is applicable. Most researchers acknowledge that they are only in a position to evaluate errors of competence and performance after undertaking a comprehensive analysis of errors.

The difference between the model of contrastive analysis and that of error analysis is that in the latter model it is not based on the comparison between the two languages, L1 and L2, but of the learners' real productions in context.

The study of systematic errors allows the transitory competence of the learner to be assessed. Their linguistic production, therefore, must be understood as a true representation of their transitional linguistic competence.

Referencias

Corder, S.P. (1974). *Error analysis. The Edinburgh Course in Applied Linguistics, vol. 3,* London: Oxford University Press, (pp. 121-142).

Chomsky, N. (1959). A review of B. F. Skinner´s Verbal Behaviour. *Language Learning*, 35, 26-58.

Fernández López, S. (1990). Análisis de errores e interlengua en el aprendizaje del español como lengua extranjera. *Tesis Doctoral.* Madrid: Universidad Complutense de Madrid.

Fernández, S. (1997). *Interlengua y análisis de errores en el aprendizaje de español como lengua extranjera*. Madrid: Edelsa Grupo Didascalia, (pp. 44-48).

Santos Gargallo, I. (1993). *Análisis contrastivo, análisis de errores e interlengua en el marco de la lingüística contrastiva*. Madrid: Síntesis, (pp.91-96).

Santos Gargallo, I. (1999). *Lingüística aplicada a la enseñanza-aprendizaje del español como lengua extranjera*. Madrid: Arco-Libros, p.30.

Santos Gargallo, I. (2004). El análisis de errores en la interlengua del hablante no nativo en Sánchez Lobato, J. and Santos Gargallo, I. (dirs.) (2004): *Vademécum para la formación de profesores: enseñar español como segunda lengua (L2)/ lengua extranjera (LE)*, (pp. 391-410).

Vázquez, G. (1991). *Análisis de errores y aprendizaje de español/lengua extranjera.* Frankfurt: Peter Lang.