

GLOSAS DIDÁCTICAS

REVISTA ELECTRÓNICA INTERNACIONAL
ISSN 1576-7809

RESPONDING TO DIFFERENT COMPOSITION TOPICS: A QUANTITATIVE ANALYSIS OF LEXICAL ERROR PRODUCTION

*María Pilar Agustín Llach / Soraya Moreno Espinosa / Almudena Fernández Fontecha
Universidad de La Rioja*

1. Review of the literature

Writing is an important skill in language, second (L2) and first (L1), development and instruction, and it furnishes relevant criteria for determining learners' language proficiency at a given point in the process of language acquisition. Writing correlates highly, actually the highest of the four language skills, with overall language proficiency (Koda, 1993; Jennings et al., 1999). Effective and fair writing assessment techniques play a critical role in student placement, evaluation of individual competence and progress, and in determining instructional needs. The design and evaluation of written tests, usually in the form of test essays, is not a trivial issue in second language instruction.

Among the many different evaluation criteria used in assessing compositions, e.g. length, discourse organisation, syntactic complexity (Cassany, 1999; Crusan, 2002; Jarvis et al., 2003) lexical measures (lexical choice, lexical originality, frequency of lexical items) are the most popular (Engber, 1995; Laufer and Nation, 1995). The direct or indirect role of lexical errors as an assessment criteria cannot, therefore, be denied (Engber, 1995). However, the use of lexical errors as a measure for determining writing quality should better be combined with other assessment criteria to achieve, thus, a most objective judgement.

Apart from devising and using reliable assessing techniques, second language teachers must be very cautious in the design of language tests. Given the repeatedly proved close (cor)relation between the written product and learners' linguistic competence, a good and reliable writing test is warrant of a valuable measure of proficiency. Unfortunately, the impact of the topic task, also called task or topic effect (Jennings et al., 1999) on the writers' response to the writing task has been a neglected variable of test design in performance analysis and writing assessment research. That the testing task has decisive effects on the

linguistic data elicited has been a long recognized issue in the field of second language acquisition (see for example Tarone and Parrish, 1988; Larsen-Freeman and Long, 1991: 31) which seems to have passed overlooked in second language writing teaching and learning research.

Nonetheless, several studies have proved the impact of composition topic in the learners' response. For example, Koda (1993) worked with American learners of Japanese as a second language and he was able to identify significant quantitative and qualitative differences between two writing tasks: narration and description. These differences demonstrated that both tasks imposed varying linguistic and rhetorical requirements. Narrative discourse proved to be more difficult, in the sense that it involved higher linguistic and rhetorical processing, and consequently obtained a lower quality score. Although not thoroughly commented by Koda, a very interesting aspect appears in his data: a high positive correlation can be observed between "good vocabulary" and all other composition quality components (comprehensibility, organization, cohesion, grammaticality, sentence structure, content strength and interest value) and overall quality rating. This data is indicative of the importance of lexical measures in quality writing.

By contrast, a later examination of topic effect in testing performance carried out by Jennings et al. (1999) yielded contradictory results. The comparison of the control i.e. no possibility of choosing the topic of their topic-based examination, and experimental group, i.e. with the chance to select the topic of their test, resulted in no statistical differences in the learners' testing performance. However, while not statistically significant, the mean scores for the choice group were slightly higher than the mean scores for the no-choice group. Authors contend that, although topic effect can be felt in general language learning and performance, it is of little or no relevance in testing using a topic-based examination. Factors such as prior knowledge of the topic, the test-taker's interest in the topic, the perceived relevance of the topic or the test-taker's opinions concerning the topic proved to have no significant effect on the test-taker's performance, probably, authors claim, because differences regarding these factors were ruled by the design of the topic-based examination.

Reid (1990) also set to investigate the linguistic and rhetoric differences in learners' responses to varying topic types and topic tasks. This author concentrated on three measures: fluency or number of words, syntactic variables and lexical variables. Concerning the main focus of interest of the present study, i.e. learner's lexical performance, Reid's study offered very valuable results, since it was proved that topic type is relevant as regards the lexical production of ESL. The vocabulary chosen by the participants in their essays varied significantly with topic type. He found considerable differences in vocabulary regarding three lexical variables: average length of words, percentage of content words, and use of pronouns, and depending not only on topic task, but also on topic task and language background.

Fluency also differed from one topic task and topic type to another, nevertheless syntax was found to be invariable in all the cases.

Similar observations regarding the variation of vocabulary choice in learners' essays depending on topic type were made by Linnarud (1986). Different vocabulary content was obtained from two different writing tasks: exposition and narration. Her results serve as support for the general tendency commented here that varying writing tasks involve different linguistic competencies and elicit different language data, especially vocabulary.

Some further evidence on this point comes from the quantitative error analysis study implemented by Jiménez Catalán (1992). In the four writing tasks she analyzed in her research, variation in the production of error type in quantitative and qualitative measures is absolutely evident. These results led her to the conclusion that the particular characteristics, and idiosyncrasies of each writing task somewhat determine the type of errors to appear in students' responses.

It is reasonable, therefore, to expect that tasks vary in the demands that they impose on learners' vocabulary and consequently, the type of task will determine the lexical error types to appear in participants' written essays. With these considerations in mind, the present paper intends to find out if different writing tasks originate varying lexical errors as regards their type and the proportions in which they manifest in each particular composition task. Consistently with previous research, the design of this study included two writing tasks of the narration and description type. In an attempt to better understand the process of L2 writing and the lexical variation observed in students' responses, the present study examines the ways in which different writing tasks influence the quality and quantity of foreign language lexical performance, i.e. production of lexical errors in composition, and what consequences this lexical error variation may have on second language instruction and writing assessment.

This study was designed to explore the following research questions:

1. Is there any quantitative difference in the production of lexical errors as regards response to different composition topics?
2. Can any quantitative variation in the type of lexical errors produced be observed as regards response to different composition topics?

2. Method

Subjects

A total of 77 subjects participated in the study. These were 10 year old Spanish speaking learners of English as a foreign language. Data was collected in a primary school in Logroño, La Rioja, Spain in spring 2004. All participants were 4th graders and could be classified as early beginners of English. English is a compulsory school subject in La Rioja from the first grade of primary education onwards. According to this, subjects had received

some three and a half years of instruction in English, which makes up for a total of around 360 hours. In order to determine objectively the level of proficiency of the informants in the foreign language, two further tests, apart from the time of instruction, were performed that revealed that subjects, in fact, were low proficiency ESL learners (see below materials section).

Three intact 4th grade classes of a unique primary school rather than volunteers were chosen to participate in this study. Of the 79 pupils that constituted the three classes, 2 could not be present the day the data were collected. Furthermore, not all participants implemented all tasks. Thus, in the first composition task there is a total of 71 compositions, since 7 subjects did not take part in the first composition data collection session, and 1 further composition had to be excluded for being illegible. For the second composition task, 6 subjects were absent the day of data collection, other 5 did not comply with the instructions and wrote their compositions entirely in Spanish, and still another one had to be discarded for being illegible. Summing up, for the first composition tasks a total of 71 subjects performed correctly, and for the second 67 compositions are at the researcher's disposal for their analysis. Despite the fact that there is an unequal number of subjects for each composing task, the study is possible because the analysis are performed with the means.

All subjects reported Spanish to be their mother tongue and all, except for one girl who acknowledged speaking Chinese, claimed having no knowledge of any other language.

Materials

The instruments used for the investigation were the compositions produced by the subjects of the study. As hinted at above, two different composition topics were used for the study. A broad subject-base theme was chosen for the first composition, henceforth C1, so that the content was the least constrained by thematical and knowledge limitations. The learners had to write a letter to an English prospect host-family, where they introduced themselves and talked about their family, their school, their home town, their hobbies and interests and any other aspect of their life and liking they might deem interesting for the receiving family to know (see Appendix A). This theme was also selected, because it did not impose any constraints on the type of language, vocabulary and grammatical structures expected, and it left free way to the learners to use their imagination and employ a wide range and variety of words and structures. Subjects had 30 minutes to complete the task and no limitation was imposed on them as regards their writing but for the topic in the instructions. Nonetheless, the free character of the task allowed students to deploy all their linguistic knowledge in English.

For the second composition task, henceforth C2, subjects also had 30 minutes. They were handed out a sheet with a series of six cartoon-type pictures that outlined a story. Learners were asked to order those pictures, tell the story, and give it a title (see Appendix B). It should be noted that these pictures specified not only the content but also the rhetorical style

of the learners' compositions. That is to say, the informants were being asked to write a narrative rather than the kind of descriptive essays that were elicited by C1. Here, content was constrained by both theme and also world knowledge limitations. Furthermore, the topic of composition imposed linguistic and, above all lexical constraints, so that the range of vocabulary expected from the participants was reduced. However, compositions were not judged by their faithfulness to the original story. With these topics, it was guaranteed that the subjects would have something to write about, and differences in the resulting essays as regards content and length due to different subject knowledge were ruled out. Thus, just lexical competence played a role in differentiating learners' written performance. The resulting essays are variable in length, content, linguistic structures, and lexical items, but all respond to the instructions given in the participants' mother tongue, Spanish. These essay tasks were part of a larger battery of tests that was implemented along four weeks (one hour every week) during class time.

A cloze procedure test and a reading comprehension test were also used in this study to determine the proficiency level of the ESL learners. Analysis of these two level tests revealed that, in fact, participants could be classified as early beginners. Both proficiency level tests were taken from the *Key English Test 1*, Cambridge University Press, 2003. The data here used are part of the data collected for the research project BFF 2003- 04009- C02- 02 funded by the Spanish Ministry of Science and Technology.

Procedures

Compositions were read twice and analyzed for lexical errors. To spot lexical errors, it is necessary that a working definition of the object of study, i.e. the lexical error is disposed of. The notion of lexical competence, that is, of what it means knowing a word lies on the basis of the definition of "lexical error". With this consideration in mind, here a lexical error is defined as the deviation from the lexical norm as it appears in dictionaries and grammars. Here, the lexical norm is established by the monolingual dictionary of English *Collins Cobuild*. Every independent meaningful unit is regarded as a lexical item. Considering this, lexical errors were here strictly counted and all lexical deviations, slight and small as they may have been, were registered as lexical errors.

Once the lexical errors were identified, a linguistic description (analysis) was attempted at that would reveal the basic underlying nature of those lexical errors. Two main structural patterns were distinguished: spelling errors, e.g. *mather*, *verday*, *sarperner*; and errors in word choice derived from mother tongue influence/interference, e.g. *coleg*, *alums*, *whale asesin*. Spelling errors can be defined, basically, as errors in the form of the lexical item, giving rise either to some non-existent word, e.g. *smooll*, *tenniss*, *ticher*; or either to an existent word but with a different meaning to the one intended, e.g. *hose* for *house*, *parrots* for *carrots*. These are usually caused by difficulties to cope with the "complicated" English orthographic system.

The disagreement in English between pronunciation and orthography is also to be made responsible for a great number of spelling errors.

Errors in word choice, on the other hand, consist in using a wrong word instead of another correct one. The resulting error is an already existing word in the target language, e.g. *fathers* for *parents*, or in the mother tongue, e.g. *arroz* for *rice*, or either an easily recognizable word in the L1, i.e. an anglicized word, e.g. *lentigues* for *lentils*. These lexical errors are mainly due to influence of the mother tongue to which the majority of them can be traced back. Thus, lexical errors were classified taking into account these two main tendencies.

3. Results

The first research question asked about whether there was any difference in the production of lexical errors according to different composition tasks, in other words it asked whether production of lexical errors would increase, decrease, or stay still in the second composition task with regard to the first. The analysis of the data revealed a somewhat controversial, striking, and non-expected result, since in absolute frequencies, subjects committed more lexical errors in the first than in the second composition task. A total of 639 lexical errors was obtained in the analysis of C1 which contrasts with the 495 lexical errors in C2. The mean of lexical errors per composition was also higher in C1: 9 lexical errors in each C1 composition and 7,38 lexical errors every C2 composition. These results are summarised in Table 1.

TABLE 1. Summary of the tallies of lexical errors per composition

	Composition 1	Composition 2
Lexical errors	639	495
Nº compositions	71	67
Mean Nº lexical errors/composition	9	7,38

This finding contradicted the expected result taking into account the more linguistically demanding character of the second composition task, that imposed a higher level of difficulty on the participants. Searching for a possible explanation for these results, the length of the compositions was analysed. This analysis turned out to be very telling and it shed light on the issue. Using the length of composition it was found out that subjects wrote less than half so much in C2: 73,37 words per composition and 147,85 words per C1 composition. This again led the researcher to a very interesting result, namely that, meanwhile subjects committed a

lexical error every 16,42 words in C1, they produced a lexical error every 9,93 words in C2. This implies that C2 compositions display a higher lexical error density than C1, i.e. performance data is marked with a higher degree of deviancy in C2. Lexical errors are, therefore, more frequent in C2 than in C1. Results are summarised in table 2. This result confirms the belief that when the thematical and linguistic difficulty of the task is increased the performance of the learners as manifested in their frequent production of lexical errors decreases (Koda 1993).

TABLE 2. Mean length of composition and lexical errors

	Composition 1	Composition 2
Mean length of composition*	147,85	73,37
Words/ lexical error	16,42	9,93

*measured in words

The second research question asked whether a quantitative difference could be observed in the production of lexical error type depending on the topic of composition. Results show that, in fact, there is a considerable difference in the type of lexical errors produced by learners in each of the composing situations. As can be seen in tables 3 and 4 spelling errors account for over two thirds of the total of lexical errors, 72,93% of the total are spelling errors in C1, whereas word choice errors are the bulk of lexical errors in the second composition task, 56,36% are word choice and 43,6% are spelling errors in C2. There a reversal of around 30% in lexical error type relative to composition topics. Therefore, the second research question must be answered positively, since a quantitative difference in the production of lexical error type was found out regarding C1 and C2. This result is consistent with the findings of previous research that revealed the variable performance of learners when responding to different composition tasks (see review of the literature).

TABLE 3. Percentage of lexical error type in C1 and C2

	Composition 1	Composition 2
Spelling errors	72,93%	43,64%
Word choice errors	27,07%	56,36%

TABLE 4. Lexical error type in C1 and C2

	Composition 1	Composition 2
Spelling errors	466	216
Word choice errors	173	279
Total	639	495

4. Discussion

In brief, this study yielded the following results: participants committed lexical errors more frequently in the second composition task than they did in the first, and qualitatively (type of lexical errors) and quantitatively different composition tasks give rise to different lexical errors. The first finding answers the first research question, i.e. if the composition task changes, the performance of the learners change relative to their production of lexical errors. This result was by no means surprising, since previous research had already found out the impact of (different) topic tasks on the students' responses to the writing task (Linnarud 1986, Reid, 1990, Koda 1993). Furthermore, the fact that the first composition task proved to be easier for the participants was not striking either. It was a broad- subject topic of the description type that contrasted in its lexical and other linguistic and rhetorical demands, e.g. grammatical structures, syntactic organisation) discourse planning, with the narration task imposed by the second composition. In the same vein, Koda (1993) also found out that "narrative discourse involves more demanding linguistic processing, at varying levels, than descriptive discourse" (p. 332), i.e. narration is a more difficult task than description. And if it is more difficult, it is logical that it will give rise to more (lexical) errors.

That the second composition task was more difficult than the first is not only revealed by the higher density of lexical errors in that one, but suffice it to think for this purpose of the fact that 5 participants were unable to respond to that task in English and had to use their first language (L1= Spanish) to comply with the assignment. Further evidence in favour of the higher difficulty of the second composition task is the absolute mean length of composition measured by the number of words. Compositions in the second task are drastically reduced to less than half of the length displayed by C1s. According to previous research (Reid, 1990; Crusan, 2002), length of composition is an important factor in writing assessment and indication of quality writing, as well as a measure of *fluency*. Although a long essay does not necessarily imply a good essay, it is very often indicative of fluency and, indeed, "length of essay has correlated highly with quality writing" (Reid, 1990: 195).

In this study, presence of lexical errors appears hand in hand with length of essay, and thus, with fluency and quality writing. Topic task has, thus, relevant effects in the responses elicited not only as regards the linguistic variables present in each composition, e.g. percentage of passive structures, of complex sentences, or of pronouns just to cite a few (cf. Reid, 1990:192), but it also has important consequences in essay fluency, and quality writing, understood as number of words and percentage of lexical errors, respectively.

Although interpretation of the data is complex, it is still possible to determine that in the light of the results, subjects were more knowledgeable and comfortable with the first task. Despite the fact that both topics provided a similar amount of information, the narration task (C2) required somewhat more cognitive complexity than the more straightforward and personal description topic of C1. Informants in C2 are detached from their personal, internal world and asked to tell a new, directed story; it seems obvious that they will encounter more difficulties in this task than in talking about themselves and their nearest environment (school, family, house). From a purely lexical point of view C2 demanded from them concrete, particular vocabulary, they did not have, in most cases; whereas C1 admitted a wider range of lexical items. Summing up, the vocabulary chosen by participants in the compositions varied considerably with topic type (cf. Reid, 1990: 198) as evidenced by the analysis of lexical error type.

Regarding the second research question, the results of the present study indicate that different composition tasks elicit different types of lexical errors. This result is consistent with the previously noted fact that both topic tasks demanded different vocabulary. In C1 the overwhelming majority of lexical errors are of the spelling type, in C2, in contrast, it is word choice errors that make up for the most numerous group of lexical errors. One possible explanation for this reversal in the production of lexical error type lies in the variable cognitive and especially lexical demands of C1 and C2.

As pointed out earlier in this section, the vocabulary required in C1 belonged to the sphere of interpersonal relationships, of the family, of the self. Vocabulary, by the way, freely chosen by the participants and to a great extent mastered by them. This explains why in this compositions, spelling errors are so numerous. Although learners positively know the words they are using, the lack of agreement between the phonetic and orthographic systems of English, and the for Spanish natives difficult English spelling originate a great confusion as regards how to write a word. For example, it is a frequent case in the present data that learners write the English word as it is pronounced, e.g. *ould* for old, *biutiful* for beautiful, or *haus* for house. Some other cases, subjects omit, add or change some graph, e.g. *scool* for school, *havee* for have, or *bahtrroom* for bathroom. Nonetheless, and despite these errors, learners show some lexical knowledge of the words at stake.

By contrast, word choice errors manifested mainly in a complete language shift, i.e. insertion of the L1 word (Spanish) in the L2 syntax (English) seem to generate in a total lack of knowledge referring the lexical item targeted. When having to face the “lexical challenge” imposed by the second composition task, informants often opt for the introduction of the Spanish equivalent of the unknown word in English, e.g. *My grandmother is fumando*, (Eng. smoking), *There are the animals: hienas, tiger, ...* (Eng. hyena). On some other occasions a Spanish word is adapted to the phonology or spelling of English, e.g. *The lats, the food are on the floor*, (Sp. *latas*, Eng. tins), *This family [...] swim in the lagune* (Sp. *laguna*, Eng. lake).

In a nutshell, the different lexical demands imposed by each of the composition tasks seem the most probable and reasonable interpretation/explanation for the results that revealed different lexical error type depending on topic task.

It is hoped that the results of the analysis, cautiously interpreted, might be sufficiently concrete to have some practical application to practical teaching. The next section will briefly outline and comment on what implications these results carry for foreign/second language pedagogy.

5. Pedagogical Implications

Because of the modest sample size, implications based on the present results must be taken with caution. Still, if, as this study has shown, the nature of the writing task affects such variables as lexical error production, then, choosing fair and appropriate testing topics should be the teacher's, and especially the tester's most important and responsible activity. The limitations of the existing writing ability testing procedures and techniques have to be acknowledged by teachers and practitioners and testing items suitable to the competence being evaluated have to be designed. The differences and variation in the lexical performance of learners depending on the different writing task demands are important factors to be taken into account by the test designer when devising and deciding on the topic of the test items to be used, since these testing topics will influence the responses of the learners and, therefore, the results of the test. Designing fair, appropriate and performance-comparable topic tasks is the next challenge of foreign and second language writing teachers.

The findings of this study also suggest some general recommendations not only for writing assessment, but also for second language instruction. First, the analysis consistently demonstrates that the two writing tasks involve different lexical competencies required for successful performance. It is essential that the writing task selected for either assessing or teaching should be adjusted to the learner's proficiency level. Suffice here to remember the different lexical error types generated by both composition topics, being C2 lexical errors clearly due to a more general lack of lexical competence than C1 errors. Consequently, the second composition topic seems more appropriate for learners in a further language

developmental stage than the first composition topic. Describing one's school, family or house on the basis of personal experience, a familiar topic, requires lexical processing procedures and vocabulary resources quite different from those used in a narrative essay.

Second, teachers should teach their learners how to face the varying lexical demands posed by the different writing tasks so that their performance will be enhanced in future essay tasks. Providing learners with lexical communication strategies, such as paraphrasing or use of cognates is a good way to prepare them for writing tasks that will demand a higher lexical competence. Vocabulary exercises should also be incorporated in composition instruction to provide a linguistic scaffolding for a given writing task. A series of exercises ranging from "easy" context-detached word lists to more comprehensive vocabulary in context for prewriting activities can help enlarge the vocabulary knowledge of the learners, and thus, facilitate subsequent composition tasks by reducing lexical problems.

6. Conclusion

Data from the present research points to the fact that choice of task has an effect on the type of vocabulary that is elicited and, consequently, on the lexical errors that appear in compositions. The quantitative differences in absolute lexical error production and in lexical error type that exist between the two writing tasks suggest important quantitative and qualitative variations concerning lexical errors and task type that are consistent with previous findings.

The limited sample size of the study and of the task types analysed make that results and implications need to be derived with caution. Generalisations must be carefully made and results should not be extrapolated out of the Spanish primary school context and the English learning situation here encountered. The lexical error taxonomy here used has proved to be a valid and reliable instrument of lexical error classification.

Future research in this field should tackle the issue of lexical error type variation with comparison of the three main writing tasks considered in the literature: description, exposition or argumentation and narration. Analysis of other lexical measures like percentage of content words, of nouns, verbs, adjectives and adverbs or lexical originality and variation should be included in future research on the variation of vocabulary in relation to choice of task.

Acknowledgements

This research was supported by a grant (FPI-2004) funded by the Comunidad Autónoma de La Rioja.

7. References

- CASSANY, D. (1999) *Construir la escritura*, Barcelona, Paidós.
- Collins Cobuild English Dictionary (1995) London: Harper Collins Publishers.
- CRUSAN, D. (2002) "An Assessment of ESL Writing Placement Assessment", *Assessing Writing*, 8, 17-30.
- ENGBER, C. A. (1995) "The Relationship of Lexical Proficiency to the Quality of ESL Compositions", *Journal of Second Language Writing*, 4 (2), 139-55.
- JARVIS, S. et al. (2003) "Exploring Multiple Profiles of Highly Rated Learner Compositions", *Journal of Second Language Writing* (in press).
- JENNINGS, M., FOX, J., GRAVES, B., and SHOHAMY, E. (1999) "The test-takers' choice: an investigation of the effect of topic on language-test performance", *Language Testing*, 16 (4), 426-456.
- JIMÉNEZ, R. M. (1992) *Errores en la producción escrita del inglés y posibles factores condicionantes*, Madrid, Universidad Complutense.
- KODA, K. (1993) "Task-Induced Variability in FL Composition: Language-Specific Perspectives", *Foreign Language Annals*, 26 (3), 332-346.
- LARSEN-FREEMAN D. & LONG M. (1991) *An Introduction to Second Language Research*, London, Longman.
- LAUFER, B. and NATION, P. (1995) "Vocabulary Size: Lexical Richness in L2 Written Production", *Applied Linguistics*, 16, 307-22.
- LINNARUD, M. (1986) *Lexis in Composition: A performance Analysis of Swedish Learner's Written English*, Malmö, Sweden, Liber Förlag Malmö.
- REID, J. (1990) "Responding to different topic types: a quantitative analysis from a contrastive rhetoric perspective". In *Second Language Writing: Research Insights for the Classroom*, Kroll, B. (ed.), Cambridge, CUP, 191-210.
- TARONE, E., and PARRISH, B. (1988) "Task-Related Variation in Interlanguage: The Case of Articles", *Language Learning*, 38 (1), 21- 43.

APPENDIX A

Instructions for composition topic 1

Imagina que vas a vivir con una familia inglesa en Oxford durante un mes. La familia se llama Mr. and Mrs. Edwards y tienen dos hijos: Peter and Helen. Escribe una carta, preséntate y háblales de tu ciudad, tu colegio, tus hobbies y cualquier otra cosa interesante que desees añadir.

APPENDIX B

B.1 Instructions for composition topic 2

Aquí tienes una serie de viñetas desordenadas. Corresponden a una historia. Primero, ordena las viñetas, luego cuenta la historia y finalmente, ponle título.

B.2 Series of the six cartoon-type pictures of composition 2

