

opción

Revista de Antropología, Ciencias de la Comunicación y de la Información, Filosofía,
Lingüística y Semiótica, Problemas del Desarrollo, la Ciencia y la Tecnología

12

Año 32, diciembre 2016, N° Especial

Revista de Ciencias Humanas y Sociales

ISSN 1012-1537 / ISSN-e: 2477-9385

Depósito Legal pp 198402ZU45



Universidad del Zulia
Facultad Experimental de Ciencias
Departamento de Ciencias Humanas
Maracaibo - Venezuela

Opción, Año 32, Especial No.12 (2016): 784-793
ISSN 1012-1587 / ISSN: 2477-9385

Interlanguage or technology? Capitalization in a learner corpus of english as a foreign language¹

Milagros Torrado y Carmen Font

Universidad Internacional de La Rioja (España)

milagros.torrado@unir.net

carmen.font@unir.net

Abstract

This article analyses a specific corpus from students of English for the Degree of Early years Education at UNIR University, Spain. It collects the input of more than one hundred students in the subject “ICT tools applied to the learning of English language”. The corpus is being analyzed by a group of researchers from the English Department at UNIR, so as to offer a taxonomy of errors and mistakes in the frame of undergraduate degrees. The results examine the use of capital letters in the data collected. The inputs have been sorted out with a specific corpus management program.

Keywords: corpus linguistics; capitalization; error correction; concordance; ICT.

¹This work has been carried out in the frame of the emerging research projects “Análisis de patrones de errores gramaticales, estilísticos y de léxico en lengua inglesa a partir de producciones escritas de inmigrantes digitales”, (Universidad Internacional de La Rioja, 2014-2015).

¿Interlengua o tecnología? Uso de mayúsculas en un corpus de hablantes de inglés como lengua extranjera

Resumen

Este artículo analiza un corpus de estudiantes de inglés del grado de Educación de la Universidad Internacional de La Rioja (UNIR), España. Recoge las intervenciones de más de 129 alumnos de la asignatura Tecnologías de la Información Aplicadas al Aprendizaje de la Lengua Inglesa y está siendo analizado por un grupo de investigadores de UNIR para ofrecer una taxonomía de errores dentro del marco de los estudios de grado. El análisis de resultados refleja el uso de las mayúsculas en los datos recabados. Las intervenciones se han analizado con un programa específico de gestión de corpus.

Palabras clave: lingüística de corpus; mayúsculas; corrección de errores; concordancia; TIC.

INTRODUCTION

Concordance studies within a specific language corpus rarely pay attention to consistency in the use of capital letters, not even in mainstream literature of error correction in an ESL context (Melvin 2014; Ojwang, 2012). However, a proper use of capital letters reflects the subject's linguistic competence and accuracy. In a globalized world where English is the language of digital communication, the proper use of capital letters becomes more and more important as a carrier of meaning. At the same time, digital environments –from conventional computer interfaces to apps and instant messaging–, with their emphasis on speed and instantaneous communication, encourage a casual use of capitalization. Students working within the context of digital educational platforms usually relax their level of writing correctness and show an inconsistent or careless management of spelling in their written exchanges. This is especially the case with capital letters, to the point that many students may not even be aware of the significance of writing “English” with a capital “E”.

This article analyses a specific corpus from students of English for the Degree of Early years Education at UNIR University, Spain. This learner corpus collects the input of more than one hundred students in the subject entitled “ICT tools applied to the learning of English language”. It gathers the students’ written participation in an online-debate (‘foro’) assessed and marked by the faculty members in charge of the aforementioned subject. This learner corpus also compiled by the faculty members in charge of the subject throughout two academic years, i.e. 2013-2014 and 2014-2015, is being analyzed by a group of researchers from the English Department at UNIR so as to offer a taxonomy of errors and mistakes in the frame of undergraduate degrees. The corpus gathers 4,414 word types and 91,666 word tokens.

1. METHODS

This study examines the correct or wrong use of capital letters in the data collected in this learner corpus. The inputs of students have been sorted out with a specific corpus management program called AntConc 3.2., which has generated key words in context (KWIC) to indicate all the concordance instances and, therefore, assess the use of capital letters.

The level of students varies from B1 to C1. Since the corpus is not homogenous in its level of English competence, we might encounter a considerable percentage of grammar and spelling mistakes. In particular, the preliminary results of our analysis reveal that most of these occurrences are inconsistent. Once the quantitative search on the use of capitalization by students is described, we propose two hypotheses for their prevalence in light of these results. In this regard, the paper aims at opening up a much needed debate about the use and misuse of capitalization as indicative of linguistic competence.

The level of language competence in English in the subjects studied varies among individuals, ranging from C1 to B1, which is and intermediate and more advanced level. Therefore, the corpus is not homogeneous as far as the level of English is concerned. At such a disparity in levels, we have decided to analyze the occurrence of a specific error among the students of English as a foreign language, namely, the wrong use of capitalization in several English words that differ from Spanish in their literal, metaphorical or sociolinguistic

usage. Once the corpus management program called AntConc 3.2. has sorted out the corpus, the following results have been generated with regard to the use of capital letters and their wrong use in the words “English” an “I”. Thus, there are 266 occurrences of the word *English* in the corpus, and 39 of which appear in lower case. For example, the subject 05TOENICTA.txt writes “I am *english* teacher in a primary school (...)”. The fact that “English” is not capitalized here suggests an interference in the use of capitalization with the speakers’ native grammar, since Spanish does not capitalize names of languages or toponymical adjectives. In the following occurrences, though, the subjects 16TOENICTB.txt, 34TOENICTA.txt, 41TOENICTA.txt and 119TOENICTA.txt make a simultaneous use of the form and the incorrect form:

34TOENICTA.txt personal computer, for example in infotmatics.
In *English* class we had the same kind of informatics.

34TOENICTA.txt (...) *english* book that were specially written for an *english* beginners. These kind of products used new (...)

41TOENICTA.txt (...) Also it is very important to experience with *English*.

41TOENICTA.txt (...) if we want children to learn it step.
Undoubtedly ICT are very important in the *english* classes.

119TOENICTA.txt (...) songs, flashcards... But, maybe when she teaches *English* as a language she doesn't use these.

119TOENICTA.txt (...) as a teacher in a Primary school; the *english* teacher in my school, she uses Power point (...)

Leaving aside other types of grammar mistakes, such as concordance subject-verb “English book that were specially written”, “for an English beginners”, “these kind of products”, or wrong choice of word, such as, “informatics”, we realize that poor grammar performance is related to an uneven choice of capitalized/non capitalized use of the word “English”. Students tend to write it in capital “E” when they refer to the English language, but fail to do so when it refers to the person’s origins. Even though this can also be explained by a careless or indiscriminate choice of capitalization, the fact that a pattern –however loose– can be found reveals a conscious choice of vowel type.

The capitalization of the pronoun “I”, however, is more problematic in terms of finding a pattern that follows through any conscious decision on the part of the student. AntConc detected 1,872 occurrences of the first person singular pronoun “I”. Of these, there are only 14 occurrences of the pronoun “I” written in lower case, such as in the individual 126TOENICTA.txt in the sentence “*i* speak to them in English all as *i* can. In our class there isn't any (...)” This obvious incorrect use of the pronoun “I” forms a contrast with the correct usage of the word “English” as far as capitalization is concerned, which suggests that the student was merely careless in making the effort to push the CAP key in the computer. Or, alternatively, the student might have wrongly believed that capitalization was optional.

However, in the subjects 32TOENICTA.txt, 40TOENICTB.txt, 43TOENICTB.txt, 47TOENICTB.txt, 125TOENICTB.txt and 128TOENICTB.txt we see once again the simultaneous occurrence of the correct and incorrect usage of the capitalized first person singular pronoun:

32TOENICTA.txt (...) a method of learning. In my opinion when *I* was at university ICT was an useful element

32TOENICTA.txt (...) *i* think that digital natives find this method (...)

40TOENICTB.txt Hello everybody! As *i* said in my first participation, *I* believe is (...)

43TOENICTB.txt (...) *i* think is very important keep in mind the (...)

43TOENICTB.txt (...) decide just for one type of material, *I* agree with the two types of resources (...)

47TOENICTB.txt (...) said both of them are really useful and *i* think each one has got a moment.

47TOENICTB.txt (...) solve this in a good way. So, finally, *I* think the best we can do is use (...)

125TOENICTB.txt (...) a fantasy-character talks. And your game *I* see with my little eyes.. *i* use (...)

128TOENICTB.txt (...) help you to be more creative. So like *I* said before, the balance is the way.

128TOENICTB.txt (...) these material will be useful in the future. *i* saw that website before is from (...)

These occurrences display an uneven usage of capitalized and non-capitalized “I” within the same sentence, and/or in different sentences from the same individual. In one case (128), the individual does not even capitalize the pronoun after a full stop, thus revealing a careless attitude and not so much a conscious choice. However, the collocation “I think” seems prone to suffer from lower casing, which in turn may suggest that the individual capitalizes the pronoun only when it introduces verbs with a stronger semantic meaning.

2. CAPITALIZATION AND CONTRASTIVE ANALYSIS

Contrastive Interlanguage Analysis (CIA) is still a valuable method in Lerner Corpus Research when it comes to identify a variety of features distinctive of learner language and assess their widespread use within a sample population (Callies & Paquot 2015), such as the one collected by AntConc. Even though the learner corpus gathered from the forum in the subject “ICT tools applied to the learning of English language” is relatively large and still a work in progress, we can draw two general hypotheses that allow us to understand capitalization as a learner error, and not simply as a slip of the computer key.

The first one revolves around the non-compulsory nature of capitalizing both pairs (“I”/“i” and “English/english) in the native language of the subject –in this sample, all of the individuals speak Spanish as a first or second language. The other languages in contact with Spanish here are the Basque language, Galician, and Catalan, and these do not interfere with capitalization usage since they follow the same rules as Spanish in this regard. Since in Spanish capitalization is unmarked in these environments where it is compulsory in English, individuals relax their own capitalization rules, even though they can tell the grammatical difference between an adjectival and non-adjectival usage of the word “English”, which in turn may trigger a loose pattern of capitalization for the same grammatical category.

The second hypothesis concerns the use of ICT. The texts included in the corpus belong to the activity in the on-line forum with the students. This forum lacks a spell checker tool, such as the one we may find in MS Word, for instance.

If we take the first hypothesis into account, we must bear in mind as well the learning phase in which students find themselves. Since their language competence is not uniform, but varies from B1 to C1, it is to be expected to find a wide variety of errors and disparity in language skills throughout this learning curve. Thus, we observe a group of students who are not yet in command of the foreign language and, as a result, they find themselves in the learning stage of “interlanguage”, a term coined by Selinker in 1972 to refer to the linguistic competence of the subject who is learning the target language and still grapples with it.

In this interlanguage period, it is fairly common to spot errors which are the result of the transference of the native language patterns of the speaker –punctuation and syntax being two of them (Alaimid & Ahmed 2012:2; Williams 2003). Another sound example of common mistakes in an interlanguage phase is ‘false friends’, when the individual uses words whose etymological root coincides with that of the target language, even though both have evolved differently in terms of meaning for each language (*pretend* – pretender < Latin ‘praetendere’).

Bearing these factors in mind, the lack of capitalization in the first letter of the words sorted out by AntConc may be due to this transference between the native language/s of the student. In those cases where the term is used correctly and incorrectly by the same individual, we can invoke interlanguage interference that would explain the recurrence of this error. As such, it would be a totally predictable error when speaking both languages and it could even be tackled in language textbooks –such is the case, for instance with, “people is/are”. However, capitalization rules are not usually the focus of grammar books and language instruction in a L2 context, and therefore the hypothesis of language transference seems to be stronger as a framework to study patterns in capitalization errors than external factors such as an ICT interface. At the same time, the patterns of transference in an interlanguage phase remit as long as the individual improves his language skills –until they disappear completely. Capitalization errors, being perceived by speakers as ‘non grammatical errors’ may take longer to remit. Until they do, the student goes through an ambiguous phase in which he does not always remember the norm and he can make mistakes despite the fact that he knows when and how to use lower case and capital letters.

The problem with the interlanguage theory at this stage is that we can also find the same type of error in native speakers who know the rules. In the following examples:

1. “Feel free to call again as i never know when another migh become available”, extract of an email message (June 2015) from a male native English speaker, 45-55 years of age.
2. “Love it i think that has actually happened”, extract of a Facebook post (August 2015) from a female native English speaker, 25-35 years of age.

The fact that native competent speakers of English make these types of mistakes in their own language and in two commonplace words leads us to think that this may not be a classic case of interlanguage interference. It does not seem to fit in the typology of intralingual interference either, since the use of the capital letter in the first person singular pronoun does not admit exceptions in English and the syntactic obligation to include the subject in non-imperative sentences explains the very frequent use of the pronoun “I”. Therefore, despite the logic behind the previous hypothesis to understand this kind of errors, we must bear in mind another variable: the use of technology. The fact that the examples from this corpus come from an on-line forum takes us to a new level of analysis. We not only have to take into account the errors from an interlanguage viewpoint (even though these cannot be overlooked either). We must not neglect the impact of information and communication technologies (ICT).

At this point we are faced with a case of dual transference. On the one hand there is an interlanguage influence, and on the other, the influence of short messages in similar non-academic fora of instant messaging. Even though studies by Wood, Kemp and Plester (2013) demonstrate that the use of technologies, in particular instant messaging systems, do not leave an imprint in the writing quality of the individual, the case studies from our AntConc sample show that we may find an intralingual interference between the typical writing of the text messages (textisms) and the writing in more formal contexts such as a forum in an online classroom environment. The online forum is for the students an exercise that they have to submit in order to be assessed, and therefore most of them strive to write correctly.

However, even in these cases, we have to add up the spell checker factor, which is activated in many ICT systems and devices. In the case of the online forum for this subject, the student lacks a spell checker, so the writing has to be supervised by the student. Text processors such as MS Word bring with it a spell checker embedded, and this fact, coupled with the sheer speed of contemporary society, produce posts which are written more carelessly than it would have been the case if written by hand. We find ourselves at a juncture in which the borderline between the written and the spoken language has been blurred to such an extent that we may not even define this type of online written communication as “written language” but rather as “oral written language”. This type of language demands the immediacy of oral communication but meets the ‘handicap’ of the comparatively slow pace of writing. Even though an online forum does not require such as an immediate response as an instant message (such as a WhatsApp) we may encounter a transference between written and oral languages. Thus, the messages from a forum or a social network are more and more casual in their observance of grammar and spelling rules due to the influence of instant messaging, which is a sort of modern shorthand.

3. CONCLUSIONS

The pervasiveness of the new social media and information technologies is changing not only our approach to communication in formal and non-formal educational environments, but it is also adjusting our perception of language correctness even when we communicate in our native language. Corpus linguistic studies, especially when they use concordance tools such as AntConc to sort out a large sample within a ICT environment, allows us the benefits of exploring errors and error correction in L1 and L2, both from the viewpoint of linguistics and sociolinguistics. By analyzing the use of spelling mistakes, such as the capitalization of the pronoun “I” and the adjective “English” we are effectively considering language acquisition and learning with the added perspective of the ICT medium. Our analysis at this stage of the sample collected so far points at the confluence of both interlanguage transference as well as intralingual interference with the digital medium of communication. As Dan Pontefract defined, “in pervasive learning you learn at the

speed of need through formal, informal and social learning modalities” (2013:5). What learner corpora tell us when applied to samples from an ICT environment and to spelling rules, is that the speed of unlearning might also travel too fast.

REFERENCES

- AIJMER, Karin (ed.). 2009. **Corpora and Language Teaching**. Ed. John Benjamins. Amsterdam (NL)
- ALAMIN, Abdulmir & AHMED, Sawsan. 2012. Syntactical and Punctuation Errors: An Analysis of Technical Writing of University Students Science College, Taif University, KSA in **English Language Teaching**, Vol. 5. N° 5: 2-8. DOI:10.5539/elt.v5n5
- CALLIES, Marcus & PAQUOT, Magali. (2015). Learner Corpus Research: An Interdisciplinary Field on the Move in **International Journal of Learner Corpus Research**. Vol. 1, Issue 1, 1-6. DOI: 10.1075/ijlcr.1.1.00edi
- PONTEFRACT, Dan. 2013. **Flat Army: Creating a Connected and Engaged Organization**. Ed. Wiley. London (UK).
- WILLIAMS, Todd O. 2003. A Study of Native Spanish Speakers’ Writing in English for Teachers in **Inquiry**. Vol. 8. N°1. <http://www.vccaedu.org/inquiry/inquiry-spring2003/i-81-williams.html>. Retrieved 18.08.2015.
- WOOD, Clare; KEMP, Nenagh & WALDRON, Sam. 2014. Exploring the longitudinal relationships between the use of grammar in text messaging and performance on grammatical tasks in **British Journal of Developmental Psychology**. Vol.32, 415-429. DOI:10.1111/bjdp.12049



**UNIVERSIDAD
DEL ZULIA**

opción

Revista de Ciencias Humanas y Sociales

Año 32, Especial N° 12, 2016

Esta revista fue editada en formato digital por el personal de la Oficina de Publicaciones Científicas de la Facultad Experimental de Ciencias, Universidad del Zulia.
Maracaibo - Venezuela

www.luz.edu.ve

www.serbi.luz.edu.ve

produccioncientifica.luz.edu.ve