

# A corpus-based study of Anglicisms in the 21<sup>st</sup> century Spanish press

Eugenia Esperanza Núñez Nogueroles

([eugenia@ugr.es](mailto:eugenia@ugr.es))

UNIVERSIDAD DE GRANADA

## Resumen

El objetivo de este trabajo es presentar un análisis de 2.198 anglicismos que permita determinar cuáles son las palabras de origen inglés que se emplean con más frecuencia en la prensa española contemporánea y las posibles razones de este uso recurrente.

## Abstract

This piece of research aims to present an analysis of 2,198 Anglicisms in order to determine the English-induced words that are employed most frequently in the Spanish contemporary press and the possible reasons for their recurrent usage.

## Palabras clave

Anglicismos  
Estudio basado en corpus  
Prensa  
Préstamo

## Key words

Anglicisms  
Corpus-based study  
Press  
Loanword

*AnMal Electrónica* 44 (2018)  
ISSN 1697-4239

## INTRODUCTION<sup>1</sup>

The presence of Anglicisms in Spanish daily life is constantly growing. Their use is reflected in the press, with this written medium acting as a suitable source to carry out a study on loanwords in a recipient language (Casado Velarde 2015; [Vázquez-Amador and Lario-de-Oñate 2015](#); [Gerdin Salas, Fuentes Morrison, Gómez and Kotz Grabole 2014](#); [Rodríguez González 2008](#); Gómez Capuz 2004). In this study, I employ journalistic texts contained in the [Corpus de Referencia del Español Actual](#)

---

<sup>1</sup> This piece of research has been funded by a scholarship granted by the Spanish Ministry of Education, Culture and Sports, and has been carried out at the University of Granada.

[\(CREA\)](#) to find out the frequency with which loanwords of English provenance are employed in contemporary Spanish.

The present piece of research aims to reach the following objectives:

1. To identify the most frequent Anglicisms (out of a list of 2,198) in the Spanish press at the beginning of the 21<sup>st</sup> century.
2. To classify these Anglicisms according to their frequency of use in Spanish contemporary newspapers.

## LITERATURE REVIEW

A corpus is, in McEnery, Xiao and Tono's (2006: 5) words, «a collection of (1) *machine-readable* (2) *authentic* texts (including transcripts of spoken data) which is (3) *sampled* to be (4) *representative* of a particular language or language variety» (original italics). As regards corpus design and typology, [Torruella and Llisterri \(1999\)](#) provide a detailed explanation on the different kinds of corpora, their extension, purpose, etc. Interestingly, Sánchez Pérez (2005) presents a brief history of corpora studies. According to this author, the first linguistic corpus of importance was compiled at the end of the 19<sup>th</sup> century. It consisted of eleven million words which were manually collected and analysed. In the 1930s, and later on, in the 1960s as well, lexical frequency lists were in vogue. In fact, in the latter decade the Brown Corpus (one million words) was created and computationally treated.

Nevertheless, the key issue that made it possible for this kind of studies to undergo a huge development was the emergence of computers, since they allowed the automatic process of enormous amounts of data. In the 1980s, a second generation of corpus studies started with the [Cobuild project](#), headed by J. Sinclair at the University of Birmingham (aprox. seven million words in its basic format). It was followed by the [Longman/Lancaster English Language Corpus](#), a collection of over twenty million words. In more recent times, third generation corpora such as the [British National Corpus \(BNC\)](#) –a hundred million words– provide users with advanced functionalities (for instance, it has been morphologically tagged). Sánchez Pérez (2005) indicates that, thanks to Internet spread, it is possible now to compile great amounts of text cheaply and in a short period of time. Indeed, this possibility

has been taken advantage of over the last years, and many scholars refer now to the web as a ‘corpus’ ([Kilgarriff and Grefenstette 2003](#); Ringlstetter, Schulz and Mihov 2006; [Cortina-Pérez and Moreno Jaén 2009](#), among others). As [Andersen \(2011\)](#) explains, «[s]ince the turn of the millennium it has become increasingly common to develop and explore web-based corpora, aka. ‘cybercorpora’ (Renouf 2007), resulting in a growing body of corpus-based studies using the web as its prime source of data ([Kilgarriff and Grefenstette 2003](#); Hundt, Biewer and Nesselhauf 2007).»

Several authors have proposed the compilation of corpora –or the use of already existing ones– with the goal of analysing the use of Anglicisms in their texts. [Cruz Cabanillas, Mancho Barés and Tejedor Martínez \(2009\)](#) focus on the presence of English words in the touristic sector and, in order to do so, they employ the touristic texts subcorpus that is included in the corpus of Spanish specialised-fields texts<sup>2</sup> collected by the aLiLex (Análisis Lingüístico del Léxico) research group. Rodríguez González (2003) also gives some guidelines on how to compile a corpus in order to study the Anglicisms appearing in its texts and with the purpose of creating a dictionary. [Laursen and Moustén \(2015\)](#) cover the use of specialised Anglicisms belonging to the financial jargon in Spanish and in Danish. On the grounds of the findings, they conclude that their «initial impression of randomness in connection with the choice between Anglicisms and local competitors has been documented by [the] corpus method approach». In the area of sports, [Balteiro Fernández \(2011\)](#) carries out a contrastive analysis of the appearance of English forms in the DRAE (*Diccionario de la lengua española* by the RAE), the *Nuevo diccionario de anglicismos* (Rodríguez and Lillo 1997), and the CREA.

[Márquez Rojas \(2005\)](#) deserves special mention. The author undertakes an automatised analysis of the «anglicismo terminológico integral (ATI)» belonging to the computing field. In order to perform this research, Márquez Rojas declares that the texts which make up the corpus of documents analysed in her study have been processed with the computing tools of the project Corpus Tècnic del Institut Universitari de Lingüística Aplicada from the Universidad Pompeu Fabra and with the software TACT (Text Analysis Computer Tools).

---

<sup>2</sup> Connected to this corpus, the aLiLex group has also developed a database in which they are classifying the Anglicisms they find in the specialised texts included in the corpus –*vide* Cruz Cabanillas and Tejedor Martínez ([2009](#), [2012](#)).

In 2012, a volume containing a selection of papers delivered at a seminar on the occasion of the 10<sup>th</sup> International Conference of the European Society for the Study of English (ESSE), held in Turin (Italy) in 2010, was published under the title *The Anglicization of European Lexis* (Furiassi, Pulcini and Rodríguez González 2012). In this piece of research, corpus linguistics is the tool applied to the study of the presence of Anglicisms in different languages. This methodology makes it possible to work with huge amounts of texts, thus obtaining a large number of instances of words of English provenance. As a matter of fact, «[i]n the study of Anglicisms, corpora are indispensable because they offer up-to-date source material from which new Anglicisms or new meanings/senses of Anglicisms may be detected. Through corpus-based research it is possible to [...] obtain information about frequency, period of adoption, usage context and authentic examples» (Pulcini, Furiassi and Rodríguez González 2012: 18). Therefore, *The Anglicization of European Lexis* aims, among other purposes, to «compare approaches and methodologies (especially corpus-based) for assessing the lexical impact of the English language on a European scale» (Pulcini, Furiassi and Rodríguez González 2012: 1).

The chapter devoted to Anglicisms in Spanish within this volume, by Oncins-Martínez, makes use of the *Corpus Diacrónico del Español* –CORDE– (Diachronic Corpus of Spanish) and the *Corpus de Referencia del Español Actual* –CREA– (Present-Day Spanish Reference Corpus), both compiled by the Real Academia Española.<sup>3</sup> They constitute extremely useful resources for the purpose of studying the introduction of Anglicisms in the Spanish language since they fill a gap highlighted by several authors, i.e. the lack of real and reliable data «that is accurately dated and abundant enough» (Oncins-Martínez 2012: 218). Indeed, these electronic materials «offer sounder [...] ways of exploring and characterizing Anglicisms in Spanish» and they «can help us track down the occurrence of foreign usages more systematically and assess the extent of their presence in Spanish more accurately» (Oncins-Martínez 2012: 217). By using them, the author finds out that there are Spanish words nowadays employed more and more frequently with an English-induced sense than with their traditional Spanish meanings. However, these semantic Anglicisms go unnoticed by the RAE, which does not include their new senses in its Dictionary (22<sup>nd</sup> edition, 2001). Moreover, Oncins-Martínez deals with the

---

<sup>3</sup> A full description of the corpora can be seen at <http://www.rae.es/recursos/banco-de-datos/corde> and <http://www.rae.es/recursos/banco-de-datos/crea>, respectively.

calquing of several phraseological units that have been adopted from English into Spanish, by means of a corpus-based analysis (of CORDE and CREA).

Oncins-Martínez has also delved into the use of corpora to study the phenomenon of Anglicisms in previous research. The author shows «some of the advantages of and the need for using corpora for exploring and assessing Anglicisms in contemporary Spanish» (2009a: 115), demonstrating that «corpus data can give us a better informed view» (2009a: 129) of this linguistic issue. What is more, the case study dealt with in this article (the adverb *dramáticamente* and its adoption of the new sense ‘espectacularmente’ due to the influence of English *dramatically*) uses «corpus evidence to support some of the suggestions made in the major studies on Anglicisms, and also to show how the influence of an English word can contribute to changes in the meaning of a cognate form in Spanish» (2009a: 128).

[Oncins-Martínez \(2009b\)](#) uses CORDE and CREA as well as the BNC –the last one with comparative purposes–, aiming to examine semantic Anglicisms. The author presents various case studies in order to display the different ways in which Spanish words are being affected by the influence that their English cognates exert on them.

Considering the German language, Seidel systematically analyses, from a corpus-based perspective, «the quantitative as well as the qualitative usage of English loanwords in the language of the German press» (2010: 3). In order to do so, the author collects a corpus made of three issues (n. 11, 21 and 30) of *Der Spiegel* (a popular German magazine) from the years 1990, 2000 and 2010, respectively. As a result of the extensive reading of these nine copies, the author finds 7,111 occurrences (i.e. *tokens*) of Anglicisms,<sup>4</sup> which are categorised and analysed with the aim of answering several questions on frequency, meaning, and integration. Results ascertain that, among other issues, the English loans with the highest frequency of occurrence are *Film, Internet, Konzern, Computer, Partner, Manager, Star* and *TV*.

With regard to the Italian language, Laviosa (2007) makes use of comparable and parallel corpora in order to «put forward a corpus-based methodology for the study of Anglicisms in business discourse» (123) and to report «on findings that focus on the lemma *business* and have been obtained from an initial application of the proposed methodology» (123). This way, the author’s goal is to approach the analysis of Anglicisms «in cross- and inter-linguistic business communication» (124). As

---

<sup>4</sup> This data is mentioned in page 33. However, in page 84 we read that «[t]he corpus of this study comprises of not less than 7,128 Anglicisms (*tokens*)».

Laviosa states, these corpus-based methods can be useful for descriptive linguists, technical translators and translator trainees, specialised lexicographers, terminologists and also for different applications in Language for Special/Specific Purposes (LSP) learning.

Finally, it should be underlined the great development on the corpus tools employed to study Anglicisms that has been implemented in relation to the Norwegian language. In the chapter «Corpora as lexicographical basis—The case of Anglicisms in Norwegian», [Andersen \(2011\)](#) describes the work-in-progress on the *Norwegian Newspaper Corpus* project, which attempts to automatically identify and analyse new loanwords from English. Several language processing tools have been set up to cover the areas of neologism extraction, Anglicism detection, collocation analysis and frequency profiling. This large monitor corpus of Present-Day Norwegian contains texts extracted from about 25 newspapers of different kinds and a range of topics. In September 2010, it was formed by 850 million words approx. (it daily grows around 230,000 words). In his last remark, Andersen accurately explains the advantages and possibilities offered by the corpus-based approach:

Thanks to recent advances in corpus building and technology, word formation and neology can be studied in empirical quantitative detail. The monitor corpus allows us to become less dependent on our intuitions and rely on statistical facts. The corpus-based approach is a valuable supplement to traditional lexicography/terminography, which involves manual extraction of words. It does not offer the full answer as to which forms to include and which forms to leave out, but it promises a systematic and empirically based proposal of where to start looking. This will hopefully lead to a significant reduction of manual work and a radical simplification of the task of looking for the needle in the linguistic hay-stack.

Following these authors, I attempt to add a contribution to the existing body of corpus-based studies on the use of Anglicisms in the press, specifically in newspapers published in Spain at the beginning of the 21<sup>st</sup> century.

## METHODOLOGY

In this article I develop a research project based on the Anglicisms collected by Delia Rodríguez Segura in her PhD thesis (1998). I have chosen this work because, to my knowledge, its author provides us with the most comprehensive list of Anglicisms in Spanish from recent times. First, I make a selection of the words she compiles, following a series of criteria that will be stated below. Then, I perform searches in the CREA, recording a vast number of contexts in which they have actually been employed. Finally, I obtain the frequencies of use of the mentioned Anglicisms in the Spanish contemporary press.

Rodríguez Segura (1998) includes a list of Anglicisms she found in different mass media (radio, television, newspapers...) during the 1990s. She classifies them, makes reference to several dictionaries and gives a brief explanation on each loanword. With respect to examples, the author illustrates every Anglicism with two or three instances. Therefore, the contextual information she provides in relation to these foreign words in use is very limited (there is no access to the whole range of varied linguistic contexts in which each Anglicism can be employed), since the means available for researchers at the end of the 1990s did not allow for the compilation of a wider sample of real uses.

However, nowadays, the possibilities offered by corpus tools for analysing a phenomenon in a great number of texts do allow for studies in an enormous variety of different co-texts and provide us with data that makes it possible to obtain frequencies of use. As Oncins-Martínez (2012: 219) states, quoting Rodríguez González (2003: 574), data banks currently available for the Spanish language as well as other digitalised and electronic resources have become an indispensable tool for linguists and lexicographers wishing to base their study of language use on more objective grounds. Indeed, in the last two decades, advances in the fields of corpus and computational linguistics have come to facilitate the work of linguists and lexicographers in general, and those interested in the study of Anglicisms in particular.

Therefore, taking Rodríguez Segura's (1998) list as a starting point, «I decided to look for these Anglicisms in the Spanish contemporary press in order to come by with real cases in which they were employed at a recent period of time. First of all, it must be stated that I made a selection of the Anglicisms provided by Rodríguez

Segura (1998) in the Appendix I of her thesis, leaving aside the following ones» (Núñez Nogueroles 2016):<sup>5</sup>

– Lexical Anglicisms that she collected only from oral mass media (the author gives their phonetic transcription and their English spelling form or the standard Spanish one. Thus, since I am going to focus on the written press, these cases are irrelevant for my goals)

– What Rodríguez Segura (1998) calls «calco fraseológico» (phraseological calque), i.e. exclamations, interjections, adverbial and prepositive expressions and locutions, conversational formulae, idiomatic expressions, fixed phrases, etc. Pragmatic elements fall outside the scope of this study

– Paronymic semantic Anglicisms and semantic calques<sup>6</sup> as well as frequency Anglicisms

– Initialisms with no specification on what they stand for

After applying these criteria, the final number of Anglicisms to be searched for in 21<sup>st</sup> century newspapers amounts to 2,198.

With the purpose of using a representative and balanced corpus of the Spanish contemporary press, as wide as possible, and that included a great variety of copies (of national as well as regional and local newspapers, generalist and specialised ones, etc.), I have chosen the CREA, compiled by the Real Academia Española (RAE).

The CREA consists of a vast number of texts extracted from different sources and held electronically. It is freely available at [www.rae.es](http://www.rae.es) and stores one hundred and sixty million forms approximately. It covers a temporal span that goes from 1975 until 2004 and an array of oral as well as written texts produced in all the Spanish-speaking countries. The written part has been selected from books, newspapers, journals, magazines and miscellaneous sources, and its documents are classified according to the following parameters: chronological, geographical and thematic.

Due to the fact that I aim to perform searches in a limited part of this huge corpus, I have filtered the texts by:

---

<sup>5</sup> The article I published in 2016 was also based on this selection of the Anglicisms contained in Rodríguez Segura's PhD thesis. On that occasion, I carried out a classification of Spanish newspapers according to the frequency with which Anglicisms appeared in them.

<sup>6</sup> For an explanation on these terms, *vide* Pratt (1980).



- Chronology: 2001-2004
- Medium: newspapers
- Geography: Spain
- Topic: all

In relation to the chronological period established, it must be reminded that Rodríguez Segura (1998) compiled her Anglicisms database in the 1990s by identifying very few instances of each loanword. Thus, the years 2001-2004 (inclusive), which came immediately afterwards (they constitute the beginning of the following decade), are a good selection to check the real usage of these English words she collected (which ones have spread and which ones have not, frequencies, etc.), using an online corpus that allows us to look up the Anglicisms in huge amounts of texts. Furthermore, this temporal segment is relevant because it refers to the beginning (the first four years) of the 21<sup>st</sup> century. Considering the medium, the corpus includes a total number of 55 different sources.<sup>7</sup> The section of the CREA I have employed (2001-2004, Spain, newspapers, all topics) consists of 5,836,589 words.<sup>8</sup>

An important methodological distinction between the planning I have just stated and Rodríguez Segura's (1998) should be drawn. This author simply fixes the time span of her data collection process: 1992-1998. She acknowledges that, when she started her research project, she did not establish a strict schedule of the number of TV channels and radio stations nor of the time devoted to watching them or listening to them. Anglicisms recorded from these media were identified during the time that we normally dedicate to watching TV or listening to the radio. The methodological change that differentiates Rodríguez Segura's (1998) PhD thesis from the piece of research I present here has been made possible by means of the technological advances that have taken place at the beginning of the 21<sup>st</sup> century.

---

<sup>7</sup> For a classification of these 55 newspapers in terms of the frequency with which they include Anglicisms, *vide* Núñez Nogueroles (2016).

<sup>8</sup> To obtain this data, first I have searched for the total number of words forming the section «Prensa» (in CREA, Spain, 2001-2004, all topics) in the [CREA's website](#) («Consulta de la nómina»). The global figure returned amounts to 7,841,721 words. Afterwards, I have manually selected the texts appearing in «Periódicos», leaving aside those included in «Revistas».

Following a systematic and consistent procedure, all the concordances I obtain are copied into an Excel document, which comprises five columns:

- Anglicism
- CREA number of *tokens*
- Concordance
- Source
- Topic

In the first one, the Anglicisms are introduced in alphabetical order. They are repeated in every line where they appear (i.e. they are written in the first column of each line that includes a concordance of it). This way, it has been possible to order the Excel table by different columns afterwards, without losing the information about which word the line refers to.

The second one is filled either with 0 (if no results have been obtained for this Anglicism) or 1 (making reference to the concordance of the word that appears in this line). By doing so, at the end of the column we can obtain the total number of *tokens* (of Anglicisms) that have been found in the corpus of journalistic texts.

The next section is devoted to the concordances of the Anglicisms. Thanks to them, we are able to check how the foreign terms appear in real use, shedding light, therefore, on grammatical elements such as the assignment of gender or the formation of the plural, and also lexical issues like the words with which they collocate in Spanish. All these aspects undoubtedly deserve a place in a further study on the use of Anglicisms in the mentioned recipient language.

The source column provides information about the title of the newspaper in which the Anglicism is employed. By working with a reference corpus that embraces such a variety of sources (specifically, 55 different ones), it has been possible to obtain results from a representative sample of the Spanish contemporary press.

The topic section contains the data on the domain provided by the CREA. It is necessary to clarify that it does not refer to the thematic field in which the Anglicism would itself be located but to the one where the text in which it appears has been classified in the CREA. Indeed, sometimes our expectations in relation to this point are not fulfilled: the fact that an Anglicism belongs to a certain area can make us think that it will be employed only in texts that are classed as members of this topic,

and then, on some occasions, results contradict this presupposition ([Núñez Nogueroles 2017](#)).

## RESULTS AND DISCUSSION

We must take into account that it has not been possible to state whether the concordances found for an Anglicism appeared in a same document or in several ones. Consequently, it should be made clear that the frequencies claimed below refer to the individual uses of the Anglicisms (i.e. it will be measured by *tokens*), without taking into consideration whether there is a massing of occurrences in few texts or the concordances are distributed each one in one document. It does not entail any problem for the present piece of research, since this study focuses on the frequency of use of each Anglicism, irrespective of the number of texts where it is employed. On the other hand, a relevant issue that must be taken into account is the thematic fields covered in the corpus that has been analysed, since this point does affect the results obtained (a word like «fútbol» will be more frequent if we deal with sports news than if we approach texts on politics, for instance). As has been indicated above (section «Methodology»), the option «all topics» was selected in the searches performed in the CREA, so the results that these queries returned came from the 93 different thematic areas<sup>9</sup> in which texts are classified in the corpus at hand.<sup>10</sup>

Considering the analysis of the real usage of Anglicisms in Spanish, it would be interesting to know which of the loanwords compiled by Rodríguez Segura (1998) are more frequent and which are less, since in this author's work these data could not be provided due to the fact that she did not carry out a quantitative study. The first ten words of the hierarchy elaborated after ordering the concordances retrieved from the CREA are organised as follows:<sup>11</sup>

---

<sup>9</sup> In actual fact, only 86 out of the 93 different thematic fields present any Anglicism in (at least one of) their documents.

<sup>10</sup> For a detailed classification of the Anglicisms according to the aforementioned fields, *vide* [Núñez Nogueroles \(2017\)](#).

<sup>11</sup> The whole scale can be consulted in Appendix I at the end of the present paper.

ANGLICISM	N. of <i>tokens</i> <sup>12</sup>
1 club	869
2 líder	830
3 televisión	785
4 fútbol	727
5 gol	577
6 web	428
7 goles	354
8 teléfono	354
9 turismo	317
10 líderes	243

Some reasons can be suggested to explain the location of these loanwords at the upper positions of the classification. The most recurrent word is *club*, which appears in the sample of newspapers 869 times. It can be employed in many different contexts, something that facilitates its frequent usage. The second place is for *líder*, a word very commonly resorted to in fields such as politics and sports (the same happens to the tenth Anglicism, its plural). The third and the eighth ones can be considered as internationalisms, since they are present in many languages in the world (*vide* Gómez Capuz 1998: 78-83). The fourth, the fifth and the seventh belong to the area of sports, a field full of Anglicisms. The sixth one pertains to a domain where Anglicisms appear constantly: computer science. Finally, the ninth position is occupied by the name of an extremely common activity nowadays, which is, besides, one of the main sources of income in Spain.

These results show a correspondence with the findings I obtained in a previous corpus-based study on the use of Anglicisms in different thematic fields ([Núñez Nogueroles 2017](#)), where evidence proved that the areas in which a higher number of Anglicisms appear in the Spanish press (2001-2004) are *sports, music, politics, computing and cinema and video*. As a matter of fact, *club, fútbol, gol* and *goles* can be located under the label *sports*; *líder* and *líderes* can be related to the *political* sphere; *web* belongs to the area of *computing*; and *televisión* is connected to the *cinema and video* field. Therefore, when using the CREA for studying Anglicisms in the Spanish press at the beginning of the 21<sup>st</sup> century, eight out of the ten most

---

<sup>12</sup> In CREA, 2001-2004, newspapers, Spain, all topics.

frequent English loanwords are linked to four of the five thematic areas where the use of Anglicisms is more common.

Out of the 2,198 Anglicisms searched for, the CREA (Spain, 2001-2004) has returned results for 784 (*vide infra*, Appendix I). On the contrary, 1,414 loanwords have not produced any outcomes, due to the reasons stated as follows. Firstly, the CREA searching tool has given an error message as a result on the following nine occasions: *computer-to-plate*, *D.J.*, *heavy rock*, *K.O.*, *marketing one-to-one*, *ready to use*, *tic*, *tics*, *too big to fail*. Secondly, since the words *pad* and *touch* constitute a lexical unit (a compound, although sometimes written as two separate words, as in Rodríguez Segura: 1998), I only count as 1 each time these two elements appear together (actually it just happens once in the corpus). Lastly, there are 1,404 Anglicisms collected by Rodríguez Segura for which I have obtained no concordances in the CREA (2001 - 2004, Spain, press, all topics), as shown in Appendix II.

There can be multiple reasons why all these Anglicisms are not found in the corpus I have analysed. Maybe some of them had fallen into disuse by 2001, perhaps there are others that had not yet spread in 2004, there can be some *hapax legomena*, or cases created *ad hoc* for the sake of irony or parody, or for humorous purposes (*vide* [Rodríguez Medina 2004](#)). On some occasions –a case in point might be «amigable con el usuario»–, the reason behind the absence could lie in a diatopic variation (Anglicisms that can be employed in Latin-American Spanish and not in Spain). Further research is needed, though, in order to determine whether this is factually the case or not. Sometimes we can face simple spelling mistakes that are not made anywhere else. In fact, in this list there are instances that seem to be errors that Rodríguez Segura (1998) could only find once. However, maybe we have this perception because they are forms that have not succeeded. Something that gives the impression of being a mistake can become an adapted lexical Anglicism if it spreads. That is why I have looked up all these words in the CREA. As Oncins-Martínez (2012: 223) underscores, using corpora «enables us to verify if a particular form occurs more or less frequently and therefore can be considered an Anglicism and not simply a bad translation or an isolated occurrence».

It could also be interesting to focus on some Anglicisms collected by Rodríguez Segura (1998) in their original English form and for which their translation is also employed in Spanish. A search in CREA with the same parameters (i.e. Spain, 2001-2004, newspapers, all topics) reveals that there are 4 concordances for «daño

colateral» –whilst *collateral damage* returned no occurrences. Likewise, «lejano Oeste» appears once and «alta fidelidad» is found three times, whereas *far west* and *hi fi* are not present in the CREA. In the case of *self-service*, «autoservicio» is employed in one sentence, and its variant «auto-servicio» in another one. A more remarkable example is that of «disco duro», whose search retrieves 64 concordances, whilst there are no hits for *hard disk*. However, other instances do not follow the same pattern: «cazatalentos», «preguntas frecuentes» or «hamburguesa con queso» are not used in the section of the CREA at hand, just as *headhunter*, *FAQ* or *cheeseburger* are not identified either. Furthermore, there is an example that shows the opposite tendency: «sistema de estrellas» does not occur in the corpus, while *star system* is found once.

## CONCLUSIONS

By carrying out a corpus-based study when approaching the use of Anglicisms in the press, we get to know the *tokens* of each of them in the sample analysed and we are also able to classify the loanwords in terms of the frequency with which they appear. Needless to say, the results presented in this article constitute tendencies and cannot be generalised, since they are based just on the section of the CREA I have studied. Nonetheless, having chosen a corpus characterised by being a representative and balanced sample of the Spanish contemporary press (it includes 55 different sources), we have the guarantee that the outcomes achieved in the present piece of research are a reliable reflection of the time span I have dealt with (2001-2004).

Furthermore, it should be taken into consideration that, in this study, I have limited the Anglicisms to a selection of those collected by Rodríguez Segura (1998). Thus, there can be other words of English origin in the journalistic texts stored in the CREA (2001-2004, Spain, all topics) that I have not focused on. However, the total number of different loans I have searched for amounts to 2,198, which constitutes a very comprehensive sample of Anglicisms.

In relation to the findings, 15,734 *tokens* of Anglicisms have been obtained from 784 different loans (i.e. *types*), whereas there are 1,414 words of English origin that have not produced any results. Out of these 784, the most frequent ones have

proved to be *club, líder, televisión, fútbol, gol, web, goles, teléfono, turismo*, and *líderes*. This correlates, to a large extent, with the outcomes I found out in a previous study on thematic fields and Anglicisms in the Spanish press ([Núñez Nogueroles 2017](#)).

In the present piece of research I carry out a descriptive synchronic study. It opens the possibility for developing an interesting contrastive diachronic research project in which an analysis of the most frequent Anglicisms employed in CORPES XXI at a more recent period of time could be contrasted with the results obtained in the study that has just been presented.

To conclude, the use of the CREA, an electronic resource developed by the Real Academia Española, has provided with higher reliability this study on the presence of Anglicisms in the Spanish press at the beginning of the 21<sup>st</sup> century. Thanks to the mentioned corpus, it has been possible to obtain results in terms of frequency of use «which were difficult, if not impossible, to confirm before» (Oncins-Martínez 2012: 219).

#### BIBLIOGRAPHY

- G. ANDERSEN (2011), [«Corpora as Lexicographical Basis: The Case of Anglicisms in Nowegian»](#), *VARIENG-Studies in Variation, Contacts and Change in English. Methodological and Historical Dimensions of Corpus Linguistics*, 6. This chapter is not paginated.
- I. BALTEIRO FERNÁNDEZ (2011), [«A Reassessment of Traditional Lexicographical Tools in the Light of New Corpora: Sports Anglicisms in Spanish»](#), *IJES, International Journal of English Studies*, 11.2, pp. 23-52.
- M. CASADO VELARDE (2015), *La innovación léxica en el español actual*, Madrid, Síntesis.
- B. CORTINA-PÉREZ and M. MORENO JAÉN (2009), [«La integración de la semántica en la enseñanza virtual del léxico mediante la explotación de un corpus extraído de la web»](#), *Greta Journal*, 17.1-2, pp. 41-49.
- CREA = REAL ACADEMIA ESPAÑOLA, [Banco de datos \(CREA\). Corpus de referencia del español actual](#).
- I. de la CRUZ CABANILLAS, G. MANCHO BARÉS and C. TEJEDOR MARTÍNEZ (2009), [«Análisis de un corpus de textos turísticos: La incorporación, difusión e integración de los](#)

- [préstamos ingleses en los textos turísticos](#)», in *A Survey of Corpus-Based Research*, ed. P. Cantos Gómez and A. Sánchez Pérez, Murcia, Asociación Española de Lingüística de Corpus, pp. 970-988.
- I. de la CRUZ CABANILLAS and C. TEJEDOR MARTÍNEZ (2009), «[Criterios para el diseño de una base de datos de préstamos ingleses en el español peninsular contemporáneo](#)», in *A Survey of Corpus-Based Research*, Murcia, ed. P. Cantos Gómez and A. Sánchez Pérez, Asociación Española de Lingüística de Corpus, pp. 741-755.
- I. de la CRUZ CABANILLAS and C. TEJEDOR MARTÍNEZ (2012), «[Email or Correo Electrónico? Anglicisms in Spanish](#)», *Revista Española de Lingüística Aplicada*, monográfico 1, pp. 95-117.
- C. FURIASSI, V. PULCINI and F. RODRÍGUEZ GONZÁLEZ (2012), eds., *The Anglicization of European Lexis*, Amsterdam, Philadelphia, John Benjamins Publishing Company.
- C. GERDING SALAS, M. FUENTES MORRISON, L. GÓMEZ and G. KOTZ GRABOLE (2014), «[Anglicism: An Active Word-Formation Mechanism in Spanish](#)», *Colombian Applied Linguistics Journal*, 16.1, pp. 40-54.
- J. GÓMEZ CAPUZ (1998), *El préstamo lingüístico: Conceptos, problemas y métodos*, Valencia, Universidad (*Cuadernos de Filología*, Anejo XXIX).
- J. GÓMEZ CAPUZ (2004), *Préstamos del español: lengua y sociedad*, Madrid, Arco/Libros.
- A. KILGARRIFF and G. GREFENSTETTE (2003), «[Introduction to the Special Issue on the Web as Corpus](#)», *Computational Linguistics*, 29.3, pp. 333-347.
- A. L. LAURSEN and B. MOUSTEN (2015), «[Tracking Anglicisms in Domains by the Corpus-Linguistic Method. A Case Study of Financial Language in Stock Blogs and Stock Analyses](#)», *2015 IEEE International Professional Communication Conference (IPCC)*, Limerick, IEEE Xplore. This paper is not paginated.
- S. LAVIOSA (2007), «Studying Anglicisms with Comparable and Parallel Corpora», in *The Study of Language and Translation*, ed. W. Vandeweghe, S. Vandepitte and M. Van de Velde, *Belgian Journal of Linguistics*, 21, pp. 123-136.
- M. J. MÁRQUEZ ROJAS (2005), [El anglicismo terminológico integral en los textos especializados: pautas para su tratamiento automatizado](#) [tesis doctoral], Barcelona, Universitat Pompeu Fabra.
- T. MCENERY, R. XIAO and Y. TONO (2006), *Corpus-Based Language Studies: An Advanced Resource Book*, London-New York, Routledge.



- E. E. NÚÑEZ NOGUEROLES (2016), [«Anglicisms in CREA: A Quantitative Analysis in Spanish Newspapers»](#), *Language Design. Journal of Theoretical and Experimental Linguistics*, 18, pp. 215-242.
- E. E. NÚÑEZ NOGUEROLES (2017), [«The Use of Anglicisms in Various Thematic Fields: An Analysis Based on the Corpus de Referencia del Español Actual»](#), *Anglica. An International Journal of English Studies*, 26.2, pp. 123-149.
- J. L. ONCINS-MARTÍNEZ (2009a), [«Towards a Corpus-Based Analysis of Anglicisms in Spanish: A Case Study»](#), *International Journal of English Studies*, 9.3, pp. 115-132.
- J. L. ONCINS-MARTÍNEZ (2009b), [«Why to “Fill a Niche” in Spanish is not a Bad Thing Anymore: A Corpus-Based Look at Anglicisms in Spanish»](#), *Proceedings of the Corpus Linguistics Conference 2009 (CL2009)*, p. 396.
- J. L. ONCINS-MARTÍNEZ (2012), «Newly-Coined Anglicisms in Contemporary Spanish. A Corpus-Based Approach», in *The Anglicization of European Lexis*, ed. C. Furiassi, V. Pulcini and F. Rodríguez González, Amsterdam, Philadelphia, John Benjamins Publishing Company, pp. 217-238.
- C. PRATT (1980), *El anglicismo en el español peninsular contemporáneo*, Madrid, Gredos.
- V. PULCINI, C. FURIASSI and F. RODRÍGUEZ GONZÁLEZ (2012), «The Lexical Influence of English on European Languages. From Words to Phraseology», in *The Anglicization of European Lexis*, ed. C. Furiassi, V. Pulcini and F. Rodríguez González, Amsterdam, Philadelphia, John Benjamins Publishing Company, pp. 1-24.
- A. RENOUF (2007), «Corpus Development 25 years on: from Super-Corpus to Cyber-Corpus», in *Corpus linguistics 25 years on*, ed. R. Facchinetti, Amsterdam, New York, Rodopi, pp. 27-49.
- C. RINGLSTETTER, K. U. SCHULZ and S. MIHOV (2006), «Orthographic Errors in Web Pages: Toward Cleaner Web Corpora», *Computational Linguistics*, 32.3, pp. 295-340.
- F. RODRÍGUEZ GONZÁLEZ (2003), «Orientaciones generales en torno a la elaboración de un corpus de anglicismos», in *Lexicografía y Lexicología en Europa y América: Homenaje a Günter Haensch*, ed. M. T. Echenique Elizondo and J. Sánchez Méndez, Madrid, Gredos, pp. 559-575.
- F. RODRÍGUEZ GONZÁLEZ (2008), [«Variaciones grafemáticas de los anglicismos en los medios de comunicación: tendencias y factores condicionantes»](#), in *Estudios de*

*filología inglesa: homenaje a la Dra. Asunción Alba Pelayo*, ed. G. Maceda, T. and L. Alba Juez, Madrid, UNED, pp. 269-283.

- M. J. RODRÍGUEZ MEDINA (2004), [«La búsqueda de efectos humorísticos, irónicos y paródicos a través de los anglicismos»](#), *ELIA. Estudios de Lingüística Inglesa Aplicada*, 5, pp. 105-120.
- D. RODRÍGUEZ SEGURA (1998), *Pasado y presente del anglicismo. Análisis de un corpus de anglicismos actuales en distintos medios de comunicación españoles* [tesis Doctoral], Almería, Universidad.
- A. SÁNCHEZ PÉREZ (2005), «Gestión informática de grandes recopilaciones de textos y extracción de información», in *Lengua y sociedad: Investigaciones recientes en lingüística aplicada*, ed. P. A. Fuertes Olivera, Valladolid, Universidad, pp. 333-367.
- U. SEIDEL (2010), [The Usage and Integration of English Loanwords in German. A Corpus-Based Study of Anglicisms in Der Spiegel Magazine from 1990-2010](#), Master's Dissertation, Tuscaloosa, The University of Alabama.
- M. STUBBS (2007), «An Example of Frequent English Phraseology: Distributions, Structures and Functions», in *Corpus Linguistics 25 Years on*, ed. R. Facchinetti, Amsterdam, New York, Rodopi, pp. 89-105.
- J. TORRUELLA and J. LLISTERRI (1999), [«Diseño de corpus textuales y orales»](#), in *Filología e informática. Nuevas tecnologías en los estudios filológicos*, ed. J. M. Bleuca, G. Clavería, C. Sánchez and J. Torruela, Barcelona, Milenio, pp. 45-77.
- M. VÁZQUEZ-AMADOR and M. C. LARIO-DE-OÑATE (2015), [«Anglicismos en la prensa deportiva de principios del siglo XX y XXI: estudio contrastivo»](#), *Epos*, 31, pp. 359-374.

APPENDIX I. 784 ANGLICISMS FOUND IN CREA, ORDERED BY NUMBER OF *TOKENS*

N. of different Anglicisms	ANGLICISM	CREA, 2001-2004, press, Spain, all topics N. of <i>tokens</i>
1	club	869
2	líder	830
3	televisión	785
4	fútbol	727
5	gol	577
6	web	428
7	goles	354
8	teléfono	354
9	turismo	317
10	líderes	243
11	filme	237
12	rock	200
13	software	200
14	récord	196
15	OTAN	193
16	pop	187
17	bar	181
18	dólar	173
19	DVD	172
20	clubes	170
21	futbolista	162
22	penalti	153
23	CD	147
24	sida	145
25	jazz	134
26	base de datos	116
27	estrés	115
28	correo electrónico	112
29	NASA	112
30	tenis	109
31	turística	105
32	golf	104
33	turístico	103
34	internet	101
35	estándar	92
36	campus	91
37	túnel	89
38	test	87
39	gasóleo	86
40	liderazgo	86
41	marines	70
42	goleador	67
43	FBI	65
44	filmes	65
45	disco duro	64
46	liderato	62
47	PC	60
48	misiles	58

49	hip hop	56
50	robot	55
51	blues	53
52	dopaje	53
53	gays	52
54	VIH	52
55	multimedia	50
56	set	50
57	cannabis	48
58	fans	48
59	rol	44
60	boicot	43
61	mediático	42
62	privacidad	42
63	aire acondicionado	41
64	gay	41
65	rap	41
66	e-mail	40
67	liderar	40
68	CIA	39
69	boom	38
70	esquí	38
71	crack	37
72	tenista	37
73	cóctel	36
74	folk	35
75	IRA	35
76	parking	35
77	funk	34
78	goleada	33
79	ranking	33
80	misil	32
81	stand	32
82	video	32
83	videojuego	32
84	CD-ROM	31
85	chip	31
86	inusual	31
87	mitin	31
88	punk	31
89	show	31
90	soul	31
91	turista	31
92	eslogan	30
93	glamour	30
94	láser	30
95	rally	30
96	escáner	29
97	hardware	29
98	marketing	29
99	MB	29
100	USA	29
101	aerolínea	28
102	autocar	28
103	caché	28
104	cross	28

105	house	28
106	vs	28
107	light	27
108	on line	27
109	pub	27
110	clan	25
111	córner	25
112	estándares	25
113	hockey	25
114	ATP	24
115	fax	24
116	surf	24
117	bus	23
118	en vivo	23
119	interactivo	23
120	márketing	23
121	master	23
122	single	23
123	drogadicción	21
124	flash	21
125	interfaz	21
126	share	21
127	CNN	20
128	lobby	20
129	off	20
130	performance	20
131	underground	20
132	whisky	20
133	cookies	19
134	módem	19
135	pubs	19
136	senior	19
137	versus	19
138	adicción (a las drogas)	18
139	dandy	18
140	fan	18
141	marine	18
142	rock and roll	18
143	sets	18
144	suspense	18
145	autocares	17
146	folclore	17
147	reggae	17
148	GB	16
149	kitsch	16
150	out	16
151	RAM	16
152	sprint	16
153	thriller	16
154	antidopaje	15
155	béisbol	15
156	consultoría	15
157	film	15
158	junior	15
159	microsoft	15
160	on	15

161	spots	15
162	western	15
163	bits	14
164	chat	14
165	play	14
166	rugby	14
167	spot	14
168	techno	14
169	UNESCO	14
170	birdies	13
171	boxeador	13
172	compost	13
173	fashion	13
174	FM	13
175	golfista	13
176	iceberg (ser algo la punta del iceberg)	13
177	jersey	13
178	prime time	13
179	waterpolo	13
180	casting	12
181	clown	12
182	dúplex	12
183	filmar	12
184	formatear	12
185	holding	12
186	indie	12
187	performances	12
188	pick up	12
189	robótica	12
190	standard	12
191	status	12
192	swing	12
193	tabú	12
194	tiempo parcial, a	12
195	trailer	12
196	cool	11
197	country	11
198	derby	11
199	esnob	11
200	esnobismo	11
201	filmación	11
202	gags	11
203	hall	11
204	minimalista	11
205	pop-rock	11
206	puzzle	11
207	rallys	11
208	showman	11
209	speed	11
210	WWF	11
211	break	10
212	cámping	10
213	folklore	10
214	jet	10
215	made in	10
216	party	10

217	top ten	10
218	yogur	10
219	baby boom	9
220	chequeo	9
221	computadora	9
222	disquetes	9
223	ferry	9
224	folclórica	9
225	MBA	9
226	play-off	9
227	póquer	9
228	radar	9
229	radares	9
230	revólver	9
231	rifle	9
232	snowboard	9
233	stocks	9
234	trial	9
235	zoom	9
236	ABS	8
237	barbacoa	8
238	CD's	8
239	comida rápida	8
240	copyright	8
241	disco compacto	8
242	funky	8
243	gospel	8
244	heavy	8
245	hippy	8
246	hooligans	8
247	jacuzzi	8
248	look	8
249	minimalismo	8
250	póker	8
251	premier	8
252	realidad virtual	8
253	reality show	8
254	SIDA	8
255	skate	8
256	yanqui	8
257	afro	7
258	asesino en serie	7
259	beat	7
260	bermudas	7
261	bikini	7
262	bogey	7
263	búnker	7
264	CPU	7
265	feeling	7
266	flashes	7
267	http	7
268	indoor	7
269	kit	7
270	lord	7
271	magazines	7
272	manager	7

273	megabytes	7
274	overbooking	7
275	piercing	7
276	póster	7
277	power	7
278	priones	7
279	rapero	7
280	reality	7
281	ring	7
282	rockero	7
283	sauna	7
284	sexy	7
285	sir	7
286	telefilme	7
287	walkie (talkie)	7
288	zapping	7
289	cocktail	6
290	córners	6
291	disc-jockey	6
292	elepé	6
293	esquiador	6
294	gasoil	6
295	hándicap	6
296	hippies	6
297	judo	6
298	mister	6
299	pádel	6
300	round	6
301	shareware	6
302	sketches	6
303	VIP	6
304	yanquis	6
305	zombies	6
306	(en) off	5
307	aeróbic	5
308	amplificador	5
309	apartheid	5
310	banners	5
311	basket	5
312	camping gas	5
313	chut	5
314	clubs	5
315	dance	5
316	drive	5
317	drogadicto	5
318	elepés	5
319	estandarización	5
320	estresante	5
321	folclórico	5
322	gore	5
323	heavy metal	5
324	implementar	5
325	LP	5
326	magazine	5
327	match	5
328	miss	5



329	mixtura	5
330	playboy	5
331	prión	5
332	relax	5
333	remake	5
334	revival	5
335	singles	5
336	ski	5
337	talk show	5
338	top-less	5
339	videoclip	5
340	vinilo	5
341	acre	4
342	approach	4
343	bicicleta de montaña	4
344	big band	4
345	boicoteo	4
346	boxes	4
347	brit	4
348	cascos azules	4
349	clúster	4
350	cociente intelectual	4
351	comedia de situación	4
352	computador	4
353	dandi	4
354	data-mining	4
355	dopado	4
356	DV	4
357	estrellato	4
358	footing	4
359	gueto	4
360	hacker	4
361	hamburguesa	4
362	hard	4
363	hard rock	4
364	hiper	4
365	hobby	4
366	inputs	4
367	intranet	4
368	jazzístico	4
369	jazz-rock	4
370	lavavajillas	4
371	mass media	4
372	masters	4
373	merchandising	4
374	microchip	4
375	oscars	4
376	pedigrí	4
377	play back	4
378	riffs	4
379	road movie	4
380	samplers	4
381	scanner	4
382	scooter	4
383	seniors	4
384	sport	4

385	step	4
386	tabloide	4
387	tex-mex	4
388	tie break	4
389	tiempo completo, a	4
390	tories	4
391	trailers	4
392	trash	4
393	airbag	3
394	alta fidelidad	3
395	alto standing	3
396	arty	3
397	baby	3
398	best-seller	3
399	bingo	3
400	birdie	3
401	bloc	3
402	body art	3
403	bogeys	3
404	breakdance	3
405	by-pass	3
406	bypass	3
407	cásting	3
408	cátering	3
409	chutar	3
410	cowboy	3
411	disquete	3
412	donuts	3
413	doping	3
414	driver	3
415	drogas de diseño	3
416	electroshock	3
417	escúter	3
418	esquiar	3
419	establishment	3
420	estresado	3
421	fanzine	3
422	folclor	3
423	folklórico	3
424	freeware	3
425	gadgets	3
426	green	3
427	grunge	3
428	hidrospeed	3
429	hippie	3
430	indexar	3
431	intercooler	3
432	jeep	3
433	joint venture	3
434	ketchup	3
435	killer	3
436	loft	3
437	lores	3
438	lounge	3
439	lumpen	3
440	managers	3

441	mercadeo	3
442	minimal	3
443	mix	3
444	mountain bike	3
445	plug-ins	3
446	pole	3
447	puenting	3
448	punk-rock	3
449	rafting	3
450	réflex	3
451	script	3
452	sex-shop	3
453	skin	3
454	slip	3
455	soft	3
456	squash	3
457	standing	3
458	star	3
459	stock	3
460	stop	3
461	tandem	3
462	tetrabrick	3
463	tops	3
464	tormenta de ideas	3
465	tránsfer	3
466	transistor	3
467	trust	3
468	turf	3
469	village	3
470	vip	3
471	wonderbra	3
472	aerobic	2
473	aislacionista	2
474	ambient	2
475	arcade	2
476	badminton	2
477	barman	2
478	beatle	2
479	bluesman	2
480	body (el body)	2
481	bol	2
482	bonus	2
483	browser	2
484	buffer	2
485	búnkeres	2
486	business to business	2
487	cambiar el chip	2
488	cash	2
489	challenge	2
490	chárter	2
491	click	2
492	crooner	2
493	dealers	2
494	denim	2
495	discapacitado	2
496	discman	2

497	dopante	2
498	doparse	2
499	dream team	2
500	eagle	2
501	escanear	2
502	esmoquin	2
503	fair play	2
504	films	2
505	finger	2
506	fitness	2
507	folklórica	2
508	friqui	2
509	gag	2
510	gente guapa	2
511	gentleman	2
512	gigabytes	2
513	glamouroso	2
514	gong	2
515	grammy	2
516	guru	2
517	happening	2
518	high tech	2
519	hit	2
520	hobbie	2
521	hot dog	2
522	jeans	2
523	kart	2
524	know-how	2
525	láseres	2
526	lentes de contacto	2
527	lifting	2
528	minifalda	2
529	misses	2
530	New Age	2
531	number one	2
532	off the record	2
533	outputs	2
534	outsider	2
535	pin	2
536	PIN	2
537	planning	2
538	playback	2
539	plug and play	2
540	plug-in	2
541	polaroid	2
542	putter	2
543	puzzles	2
544	ragtime	2
545	rallye	2
546	rallyes	2
547	remakes	2
548	rockabilly	2
549	sandwich	2
550	saxofón	2
551	sex-symbol	2
552	short	2

553	ska	2
554	sketch	2
555	sketchs	2
556	slalom	2
557	snow	2
558	stablishment	2
559	street	2
560	stretching	2
561	supermodelo	2
562	superstar	2
563	surfista	2
564	tattoo	2
565	teenager	2
566	telemárketing	2
567	tique	2
568	tobogán	2
569	top model	2
570	transfer	2
571	tripi	2
572	twist	2
573	unisex	2
574	váteres	2
575	videoclub	2
576	walkman	2
577	water	2
578	working	2
579	yonquis	2
580	acid	1
581	aeroclub	1
582	after hours	1
583	aislacionismo	1
584	antiestrés	1
585	autostop	1
586	background	1
587	bacon	1
588	batir un récord	1
589	bistec	1
590	blue jeans	1
591	bluesy	1
592	boy	1
593	briefing	1
594	british	1
595	brothers	1
596	building	1
597	bulldozer	1
598	bulldozers	1
599	bungalows	1
600	burn out	1
601	byte	1
602	bytes	1
603	cabezas rapadas	1
604	cámpings	1
605	cárdigan	1
606	cash flow	1
607	catering	1
608	cattering	1

609	charter	1
610	cheerleader	1
611	chequear	1
612	ciberpunk	1
613	cinemascope	1
614	climax	1
615	clip	1
616	clipper	1
617	clownesco	1
618	cluster	1
619	comic	1
620	comida basura	1
621	compact	1
622	compact disc	1
623	computer	1
624	comunicacional	1
625	core	1
626	cowboys	1
627	curry	1
628	cyborg	1
629	DNA	1
630	downtown	1
631	driblar	1
632	drop	1
633	e-business	1
634	en el aire	1
635	escáners	1
636	escúteres	1
637	esnifar	1
638	esnobista	1
639	estandarizar	1
640	eye-liner	1
641	fashion victim	1
642	fast food	1
643	faxes	1
644	feedback	1
645	flirtear	1
646	flirteo	1
647	forever	1
648	freaks	1
649	free	1
650	free float	1
651	freelance	1
652	FSH	1
653	full contact	1
654	full time	1
655	funboard	1
656	futbolero	1
657	gángster	1
658	ginger ale	1
659	ginseng	1
660	gin-tonic	1
661	glam	1
662	glamuroso	1
663	GMAT	1
664	godspell	1

665	hamburguesería	1
666	happy-	1
667	happy hour	1
668	hit parade	1
669	homeless	1
670	hooligan	1
671	hub	1
672	indi	1
673	input	1
674	jam session	1
675	jazzman	1
676	jetlag	1
677	jogging	1
678	joint-venture	1
679	juniors	1
680	just-in-time	1
681	karts	1
682	kick boxing	1
683	lady	1
684	leasing	1
685	LPs	1
686	mailing	1
687	management	1
688	marca de fábrica	1
689	mecadotecnia	1
690	melting pot	1
691	mini-CD	1
692	minidisc	1
693	minigolf	1
694	mítin	1
695	music hall	1
696	new look	1
697	noise	1
698	offset	1
699	outdoor	1
700	output	1
701	pack	1
702	panties	1
703	patch	1
704	patchwork	1
705	pecé	1
706	peeling	1
707	penalty	1
708	permafrost	1
709	ping pong	1
710	pixel	1
711	play backs	1
712	play-offs	1
713	pole position	1
714	prensa amarilla	1
715	prion	1
716	punch	1
717	punkis	1
718	putts	1
719	rail	1
720	rasta	1

721	ravers	1
722	remix	1
723	removable	1
724	rent-a-car	1
725	reverb	1
726	RISC	1
727	robocop	1
728	robotizar	1
729	royal	1
730	saloon	1
731	sampleado	1
732	sampler	1
733	sandwichera	1
734	sheriff	1
735	show business	1
736	skateboard	1
737	skinheads	1
738	sky	1
739	slang	1
740	slips	1
741	slogan	1
742	smoking	1
743	speaker	1
744	spin	1
745	spoiler	1
746	spray	1
747	staff	1
748	star system	1
749	starlettes	1
750	stoniano	1
751	storyboard	1
752	stretch	1
753	strip-tease	1
754	supermanes	1
755	superwoman	1
756	surfing	1
757	tecno	1
758	tee	1
759	télex	1
760	terabyte	1
761	terminator	1
762	terminators	1
763	tetra-brik	1
764	the end	1
765	thrash	1
766	tocadiscos	1
767	toffee	1
768	toffees	1
769	touch	1
770	transformers	1
771	trekkie	1
772	trip	1
773	trip-hop	1
774	VIPs	1
775	voltio	1
776	warm-up	1



777	weekend	1
778	West End	1
779	West Village	1
780	westerns	1
781	windsurf	1
782	yonkis	1
783	yupis	1
784	yuppies	1
Total n. of tokens		15734

APPENDIX II. 1,404 ANGLICISMS COLLECTED BY RODRÍGUEZ SEGURA FOR WHICH NO  
CONCORDANCES ARE OBTAINED IN THE CREA (2001-2004, SPAIN, PRESS, ALL TOPICS)

Acces time (*sic*), accesar, ace, acid house, acid-aficionado, acondicionador de aire, acquaplanning, action painter, action painting, acuaplaning, acustic club (*sic*), Advanced Vehicle System, aerobús, affirmative language , after games , afterpunk, after-shave, aftershaves, aftersun, agility, agribusiness, agro-pop, aguaplanning, airbus, alien, all star, All Star Game, alocar, amigable con el usuario, animatronics, Anti Block System, antiapartheid, antidive, anti-doping, antifans, antistress, anti-trust, anti-zapping, anti-zapping, aparthotel, aparthoteles, aqua training, aquaplaning, assistant manager, ataches, audiotex, auditar, autocue, autocues, autoestop, autofocus, autolink, autoreverse, autorickshaw, autorickshaws, autorreverse, autotracking, AVS, baby sitter, babylonmanía, backgammon, backstreetmanía, backup, bad-lands, baffle, banana split, bar towel, barbacuá, barra inteligente, base (salto base), BASIC, básquet, basket control, basketmanía, bas-kit, basquet, basset, basset hound, baton twirler, battletech, beat 'em up, beat generation, beatlemania, beatles, beautiful, beautiful people, beauty case, beauty people , bebida inteligente, bed & breakfast, beedies, beep, beeper, behaviorismo, behaviorista, bet.seller, bicicross, bicoid, big one, big-bang, biker, bikinis, biofeedback, bioman, bioplastic, bio-spray, bip-bip, biscúter, bit, bítico, bitmap, bits generation, biutiful (*sic*), black, black power, blackjack, blazer, blazers, blend of USA, blister, blocs, block language, blocs, blood-shift, blue chip, blue jean, blue stocking, blusero, blush, bobby, bodies, body, body building, bodyboard, bodyboarding, bodybuilding, body-building, bodymilk, bodys, bodysilk, bonus poll, boock, book, booker, booking, border collie, boss, botafumeiroman, bottombra, bourbon, box and one, box office, boxear, boxístico, boy scout, bracelete, brainstorming, brake dance, brandymanía, breading, breakthrough, brik, British Dominion of Gibraltar, brit-pop, broadcast, broadcasting , broker, broker, brokerage, brother, brunch, brushing, brushings, buckminsterfullereno, buckyball, buckytubo, buddy movie, budin, building society, bulldog, bumerán, bungalow, bunker, bunkerizarse, burger, burnout, burn-out, bus local, bus-bob, buscador de senderos, business class, business school, businessman, business-model, buy-out, by pass, byroniano, cacaburguer, cache, CAD, café society, cake, californian sun, call girl, call girls, calls, camcorder, camcorders, camel, cámel, Cancellor of the Exchequer, candies, canoeing, canyoning, car audio, Car Hi-Fi, caravaning, cardigan, cardiotraining, cartridge, cash & carry, cast, casting couch, casual wear, catcher, cazatalentos, CD-Audio, CD-I, cederrrom, cederrromes, cederrón, celler, cerdo de Guinea, cerdo guineano, cermet, cermets, challenges, chance, charlestón, chart, cheap magazine, check-in, checkpoint, cheerleaders, cheeseburger, chessball, chewing gum, chief executive, chief executive officer, chipset, choped, chopped, chopped-beef, chopped-pork, chop-suey, chorus, chou, chouleaders, Chrismats [*sic*], Christmas cake, Christmas shops, Christopher lait, chunking, chute, cibacrons, ciberman, ciber-rock, ciborg, cicloscooter, CISC, city, clase business, clasicmanía, clearing, clergyman, climatronic, clinero, clip-art, clippear, clipping, clowns, CM, coach, cockney, coke, cold shoulder, collateral damage, college, colorprinting, combot, comic book, commonwealth, compacdís, Complete Set of Instructions, composite, compositing, compostable, confirming, conservativo, consulting, container, containers, cooler, coolies, cop, copia dura, cops, cornflakes, corporate rightsizing, côtel [*sic*], cottage, cottages, country rock, crak, crash, crash test, crash-sensor, crashtest, crash-test, Crew Resources Management (CRM), crol, crolista, croll, cross training, crossbar, cuáasar, culling, cult band, cup, custom, cutter, cyberjeans, cybermall, cyberpunk, D.J.'s, dandy-hippy, darling, daseinanalysis, daseinanalyst, DAT, data-link, data-warehouse, day pack, DAZ, DCC, DDC, DDD, de mood a mood, deadline, dealer, dealing, debugear, decibelio, demo, demolition man, desflashado, design, desktop, destroy, destroyer, destroyers, differently abled, digital audio tape, digital scan, digital video, digital-VHS, dinner, dinomanía, direccionar, dirty chic, dirty protets (*sic*), dirty realism, disablear, disco digital, disco music, discount, diskette, diskettera, diskettes, display, disquetera, distress, dithering,

diving, donkey boys, donut, dopar, dopping, double face, downshifting, DPI, drag queen, Drag Racing, dragmanía, dragsters, dribbling, dries, drill, drills, drink, drops, drugstore, DSP, duffel coat, dummies, dummy, dumper, dumpin, DVD-ROM, D-VHS, eagles, earthshoes, easy care, easy mop, easy-rider, easyriders, e-cash, eco-pack, ecu, ECU, ecus, ECUS, edutainment, eggosaurio, egotismo, eigenface, electricbike, electrochoque, e-money, emoticons, enablear, encuestas de salida, equity retreat, escaner, escaners, escay, eslipada, esloganes, esmokin, espídico, espónsor, esponsorización, esponsorizar, esquí-golf, esquinas parlantes, essential Egypt, estéreo, estresada, estresor, estripista, eurobasket, eurobuilding, euromárketing, eustress, exit, expel, extropianos, extropians, eye-fate, eye-linner, factoring, fanes, FAQ, far west, fashion victims, fast, fastforward, faxear, fed funds, feelings, fender, fifty-fifty, filmlets, final cut-off time, final four, first termic, five o'clock tea, five step, flashazo, flashback, flasheado, flashear, flatlock, flavor, fleemarkets, flirt, floppy, floppy disk, folclorista, folkie, folky, follow-up, footbag, for sale, Foreign Office [sic], forfait, forfaits, formula writers, forward, fourballs, foursomes, foxterrier, fox-trot, FPS, franchising, free ball, free cinema, free enterprise economy, free lance, free lance killer, free rider, free-bol, freedom, free-lance, freetos, french cachette, fresbee, friends, fruit cake, full duplex, fullereno, fun bike, fundraising, fun-fly, funkier, funny car, funware, fútbol, fut-voley, gai, gais, galaxy rock, gangster, gangsteriles, gangsterismo, gangsters, gángsters, gánster, gánsteres, gánsters, gap, garage rock, gas oil, gas-oil, gated communities, gel-body, General Management Admissions Test, general store, gentry, Geographical Magachin [sic], ghetto, gigabyte, gigaoctet, gigaoctets, gimlet, gin-tonics, girl power, glam rock, glamrock, glam-rock, globe trotters, golaverage, golden retriever, golden retriever, golden share, goremanía, goretex, green belt, green machine, grill, grind, grogui, groupie, groupware, guest star, guonderbrá, gym-jazz, halloween, halter, hamburguesera, hammer, handicapado, handling, happy coats, happy end, happy meal, happy sixties, hard disc, hard discount, hard disk, hard-core, harris, HDD, headhunter, heli-esquí, heli-skiing, helpware, hi fi, hidrobob, hidrofoil, high school, high society, high speed, high tec, high yielder, highware, hill billy, hip hop swing, hipermedia, hipers, hippiesco, hippioso, hippismo, hippizante, hobbies, hobbys, HOG, hollywoodesco, hollywoodiano, hollywoodiense, home run, homeland, homepage, hooligang, hora feliz, hot line, hot spots, hot-ball, house generation, hovercraft, HTML, husky, hustler, imprinting, imput, inputs, in (estar in), in/out, infoshow, inicializar, inner city, intercity, interesoteniante, interface, interfase, interviú, Ironman, item, jacksonmanía, jeanswear, jerséi, jerséis, jerseises [sic], jerseys, jet boat, jet set, jet society, jet-foil, jingle, JIT, jointventure, joystick, joysticks, jr, jr., Juanas Nadie, jumbo, karting, key symbol, KeyCard, killer kiss, kilobit, kilobyte, kind start, kindergarten, king size, kitch, knowbot, ladies, laight (i.e. light), láit (i.e. light), lambswol, lambswool, land of memory, larger than life, láser disc, lásercom, lasérico, laser-writer, latin lover, LBE, lean management, leggings, leggins, LEM, lexical, life line, linkar, lip-fix, lipofiling, lipstick, lising, live, lobbista, lobbying, lobbysmo, local bus, looping, lúmpen, lumperío, lumperizarse, lunch, lunchable, lurex, machoman, mad max, madmaxista, magazin, mail art, mailart, mail-art, mailear, mailings, major, majors, management buy-out, managing director, mapear, máquina house, marching band, market markers, marketing mix, Mars Environmental Survey, master in business administrastion [sic], Mbyte, Mbytes, MD, MD walkman, MD walkmans, MDs, medical research manager, mega-bit, megabite, megabyte, megacarrier, megaoctet, mega-pack, mega-star, megastore, megatón, megatones, megatop, memorabilia, mercado de pulgas, merchandasing, merchandiser, MESUR, metal rock, metal sound, microcyborg, microfalda, micropeeling, middle class, mini cooper, minibang, minibasket, minicar, mini-casting, minidisks, mini-Lp, minimarket, minimarkets, minishort, minstrels, MIR, mites, mitin, mitines, mobile computing, modem, monis, monises, monster movies, mopa, morphing, motorball, mountain biker, mouse, mouses, MPC business audio, multimedíatico, multivan, music television, must, muteo, N.A., nailon, naked, nanny, NC, neo-hippie, neo-hippies, neo-hippy, neo-hyppy, NetPC, netting, network computer, network computing, never, new journalism, new peeling, new wave, news, newsgroup, nickname, niger, nigger, niñóbic, no-clipping, noise-pop, non food, non oil business, nonsense, nonsensuales, non-stop, non-tracking, non-vegetarian,

noqueador, noqueadores, noqueo, notebook, nurse, nursería, nylon, offshore, offshore administration, offside, off-sider, oil free, OK, okei, okei Makei, on site, on the right track, on the road, one design, op-art, operator, optical, opting-out, orsaid, orsay, óscares, outing, overbookings, overland, overline, oversize, pacekeepers, pack ergo, packaging, packansing, packing, paddel, padding, paddings, paddle tennis, padel, padle, package show, palo asesino, panty, pantys, papeles de identidad, parapenting, parka, part time, party line, passing, passwords, patcheado, patches, path, Patient Controlled Anesthesia, pay per view, PC Card, PC exchange, PC master, PC tools, PCA, PCC, PC's, peach, peanut, peanuts, pearcing, peep show, peeping show, pen, penny magazine, Pepsi board, perkins, Personal Communication Computer, personal computer, peterpanismo, petting, phone-box, picú, pinball, pin'ups, pipe-line, pit bull, planking, play-back, playbacks, play-backs, plotear, Plug&Play, plum cake, plussing, polaroids, poll (fracaso poll), polving, pompom girls, poney, ponies, pony, pop corns, pop tarts, pop-art, poper, Poppy Day, portable, portafolio de productos, poscript, posing strap, post minimal, postal free, poster, pound cake, power CD, power ranger, power trío, powerpop, preliminary cut-off time, pressing-boxeo, pressing-catch, pre-task, pretty, price smash, printable, printer, procedural, product manager, product manager senior, product placementm, progroms, prom, proms, pro-stock, pro-tory, psycho killer, psychokiller, psychokiller, pubes, Publish or Perish, pudin, pullover, pullovers, punkie, punkys, putt, quads, quark, quarks, quásar, quest, quickwheel, racing, rad board, RADS, ramie, ranger, rapear, rapper, rappero, rappers, raps, rash, rating, ratmusket, RDS, read and return, ready for use, reaggea, realiti, realiti show, realiti-chou, realitichous, realitis-chous, recordman, red localtalk, Reduced Set of Instructions, reggae-rap, remasterizado, remasterizar, remember, render, rendering, renderizado, renderizar, renting, repóker, reporting, resetear, resort, restiling, restyling, retrollamada, reverv, rewind, RFC, rhithm and blues, rhythm & blues, rhythm n' blues, richshaw, richshaws, rickshaw, rickshaws, riesgohabiente, rings, roadster, roast beef, rock indie, rocker, rockerizarse, rockers, rocketeers, rocketters, rock-soul, rock-star, rock-stars, rollerski, roll-on, roll-over, roof-garden, roquero, rosbif, rotary, rotoscoping, rough, round robins, rounds, royals, royalty, running, rush, rythm and blues, sales representative, saltobasista, samplear, sampling, sand wedge, sandwichería, sandwiches, SAR, scones for tea, scoop, scope, scrambled, scrapie, script girl, scroll, season, selfmade man, self-service, sensory vortex, serial killer, servicing a target, set point, setter, seventy's, sex-appeal, sexes, shakespeareanas, shapeCD, shares, sherpas, shifting, shilling pamphlets, shimmy, shocking, shopping, shorts, short-short story, short-story, show room, show woman, showbiz, showgirl, showroom, show-room, showtime, showwoman, show-woman, shuttle, sidafobia, sidecar, sidecares, sidecars, sides, sidoso, sillabus, sillonbol, síndrome del quemado, sistema de estrellas, sitcom, skai, skatalítico, skay, skeet, ski doo, ski kart, ski karting, skidoo, ski-doo, skikart, ski-kart, skikarting, ski-karting, skin abraser, skinhead, skinners, sky surfing, skyline, slapstick, slice of life, slip-bra, slot, slowcore, slowly wilderness, slum, SM, SMA, smart card, smart drink, smart drinks, smasheador, snack, sneakers, snipe, snipers, snipes, snow cycling, snuff, snuff movie, snuffmovie, snuff-movie, soccer, Soft Damp, soft discount, softball, soul music, soundblaster, soundtrack, speaker's corner, speakers-corner, special garments, specke, speech, speed wagon, spider, spilberiana, SPLA, splash, spoils system, sponsor, sport-chic, sportmanía, sportswear, sportwear, spread, springer spaniel, square, SRAM, stake, stake holder, stand-by, star team, starlet, starter, Static Random Access Memory, station wagon, status symbol, steadycam, stereo, stick, stockholder, storage, store, strech, strecht, streep poker, streep-tease, stretch, stresante, string, strip, stripper, subnotebook, suburbia, sudden fiction, sueter, sueters, super-bike, superfan, superfreak, superlumpen, superman, supermarket, supermarkets, superposter, super-pretty, súper-sport, super-top, supervan, supervans, surfer, surfersónicos, SurfMan, surround, swap, swaps, Swatch car, syllabus, tandoori, tap dance, task force, taylor, team manager, techno-triller, tecno-trance, teddy boy, teddy boys, teenage hits, teenagers, teleprinter, tele-working, tener un chip puesto, terabytes, territory manager, testing, tetra brik, tetrabra, tetrabrik, Thanksgiving (Day) [sic], Thanksgiving (Day), thatcher, thatcheriano, thatcherista, the American way of life, the

mail box, the prevailing wind, thrash metal, timeshifting, timing, tiquet, tombstones, tommies, tonner (sic), too much, top drivers, top fuel, top models, top secret, top-forty, toppings, topss, tory, total look, tourist liquor permit, toymanía, trabbies, traby, trackball, tracking, trade marketing, trade-off, trail, trainers, trainspotter, trashtalking, treki, trekies, treking, trekkies, trekking, trench coat, trenka, tripis, trusts, T-shirt, T-shirts, tumbing, tuner, Turbo Man, turnover, tweed, twin cam, twin set, twin sets, twinset, twin-set, twinsets, twin-sets, U.S.A., UFO, ufología, ufológico, ufólogo, ultimate fighting, ultra-dry, union jack, Unites States (sic), unpleasant, unplugged, unplugged, uplifting, vard-sales, vatio, vegetarian, vending, Very Important Person, Very Important Persons, VESA, videocasting, videogame, videowall, videowalls, vips, visiting a site, visual merchandising, vivation, voley-playa, VRC, VRLM, vueling, walkmans, wanted, washback, WASP, wáteres, waterproof, wavetable, WebCam, webcasting, web-sites, weekender, welfare, West, wets, wheeling and dealing, whiskería, whiskey, whisky on the rocks, white collar, wind surf, wind-surf, windsurfing, windsurfista, winer (sic), wiskey, wisqueria, wonderbragas, wondernalgas, woolmark, work, workflow, workshop, World Wide Web, WWW, yarda, yé yé, yes, yonki, yonkie, yonkies, yonqui, young urbans professionals (sic), yupi, yuppie, yuppillage, yuppismo, zamp, zamping, zapeo, zapineador, zapineo, zaping, zapirroides, zapista, zipper, zombie, zona dadora