

What can CORPES XXI tell us about the thematic distribution of Anglicisms in the Spanish press?

EUGENIA ESPERANZA NÚÑEZ NOGUEROLES

Universidad de Granada
eugenia@ugr.es

INDEX

1. Introduction
2. The study
3. Results
4. Discussion
5. Conclusion
6. References

1. Introduction


- Study of Anglicisms in different recipient languages: long tradition
- Recently: corpus-based perspective
- Laviosa (2007): English loanwords in Italian business discourse
- Seidel (2010): Anglicisms in the language of the German press
- Andersen (2005, 2011, 2012a and 2012b): English loanwords in the *Norwegian Newspaper Corpus*

1. Introduction


- *The Anglicization of European Lexis* (Furiassi et al., 2012)
 - **Anglicisms in Spanish:** “Newly-coined Anglicisms in contemporary Spanish. A corpus-based approach” (Oncins-Martínez, 2012)

Corpus Diacrónico del Español (CORDE)	}	Useful resources. Real and reliable data
Corpus de Referencia del Español Actual (CREA)	}	Frequency analyses
- “Towards a Corpus-Based Analysis of Anglicisms in Spanish: A Case Study” (Oncins-Martínez, 2009a)
- “Why to 'fill a niche' in Spanish is not a bad thing anymore: A corpus-based look at Anglicisms in Spanish” (Oncins-Martínez, 2009b)

2. The study

- BACKGROUND: Núñez Nogueroles (2017), PhD dissertation
 - ❖ English loanwords in the Spanish press
 - ❖ A selection of 2,198 Anglicisms from Rodríguez Segura (1998)
 - ❖ Queries in CREA: press, 2001-2004, all thematic fields, geographically restricted to Spain
 - ❖ 10 most frequent Anglicisms:
 1. club
 2. líder
 3. televisión
 4. fútbol
 5. gol
 6. web
 7. goles
 8. teléfono
 9. turismo
 10. líderes
- 

2. The study

- OBJECTIVE: to explore the thematic distribution of these 10 loanwords in the 21st century
 - in relation to their number of occurrences
 - in relation to the number of documents where they are employed
- 

2. The study

- Corpus del Español del Siglo XXI (CORPES XXI): last reference corpus by the RAE.

Time span: 2001-2015.

25 million forms per year.

Distribution:

- ☐ Geography: Latin America (70%), Spain (30%)
- ☐ Mode: written (90%), spoken (10%)
- ☐ Medium: books (40%), periodical publications (40%), Internet (7,5%), miscellany (2,5%)

2. The study

- Corpus del Español del Siglo XXI (CORPES XXI): last reference corpus by the RAE.

Bloque	Soporte	Géneros (Ficción) Temas (No ficción)	Tipo de texto
Ficción	Libro	Guion Novela Relato Teatro	Ficción
No ficción	Libro	Actualidad, ocio y vida cotidiana Artes, cultura y espectáculos Ciencias sociales, creencias y pensamiento Ciencias y tecnología Política, economía y justicia Salud	Académico Biografía, memoria Divulgación Jurídico-administrativo Libro de texto
	Prensa	Actualidad, ocio y vida cotidiana Artes, cultura y espectáculos Ciencias sociales, creencias y pensamiento Ciencias y tecnología Política, economía y justicia Salud	Académico Carta al director Crítica Crónica Divulgación Editorial Entrevista Noticia Opinión Reportaje Varios

www.rae.es

2. The study

- Corpus del Español del Siglo XXI (CORPES XXI): last reference corpus by the RAE.

Bloque	Soporte	Géneros (Ficción) Temas (No ficción)	Tipo de texto
Ficción	Libro	Guion Novela Relato Teatro	Ficción
No ficción	Libro	Actualidad, ocio y vida cotidiana Artes, cultura y espectáculos Ciencias sociales, creencias y pensamiento Ciencias y tecnología Política, economía y justicia Salud	Académico Biografía, memoria Divulgación Jurídico-administrativo Libro de texto
	Prensa	Actualidad, ocio y vida cotidiana Artes, cultura y espectáculos Ciencias sociales, creencias y pensamiento Ciencias y tecnología Política, economía y justicia Salud	Académico Carta al director Crítica Crónica Divulgación Editorial Entrevista Noticia Opinión Reportaje Varios

www.rae.es

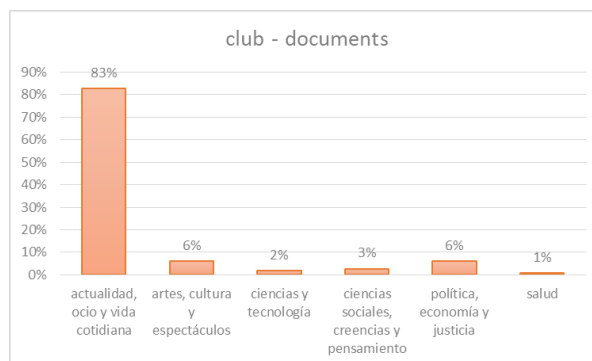
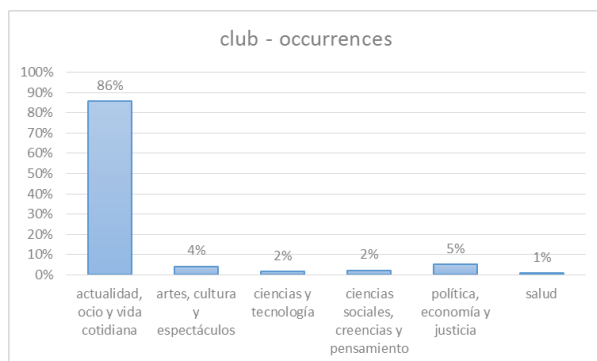
2. The study

- Corpus del Español del Siglo XXI (CORPES XXI): 2001 - 2015

Search criteria:

- Geographic area: Spain
- Mode: written
- Section: non-fiction
- Medium: press
- Thematic area: all
- Text typology: all

3. Results

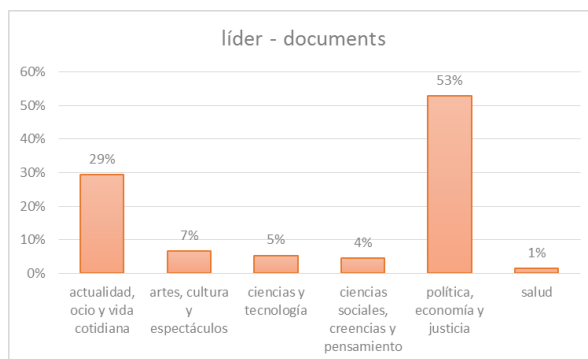
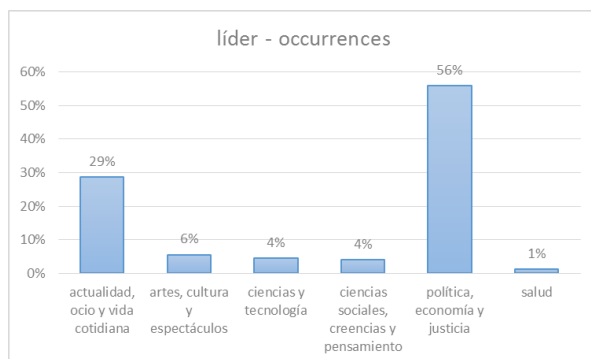


Absolute frequency: 5,381

N. of documents: 2,685

Normalised frequency: 21.75 hits per million words

3. Results

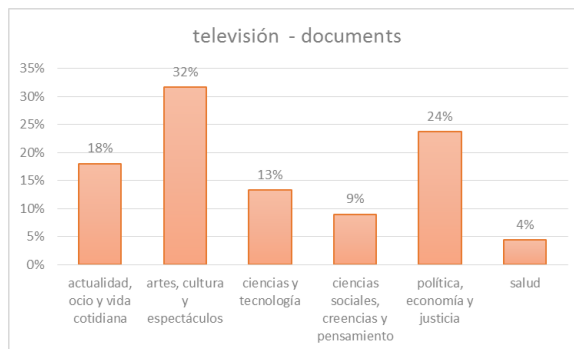
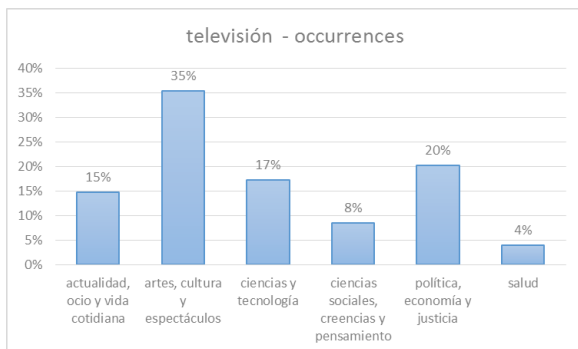


Absolute frequency: 5,776

N. of documents: 4,091

Normalised frequency: 23.35 hits per million words

3. Results

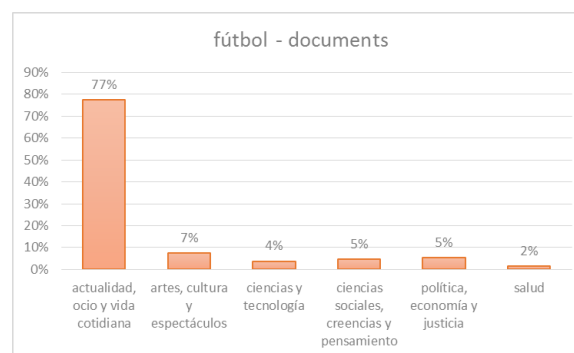
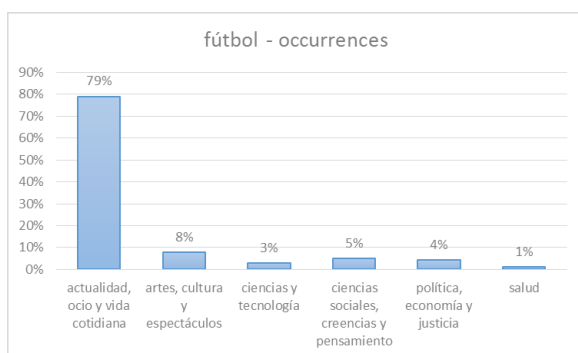


Absolute frequency: 5,808

N. of documents: 3,690

Normalised frequency: 23.48 hits per million words

3. Results

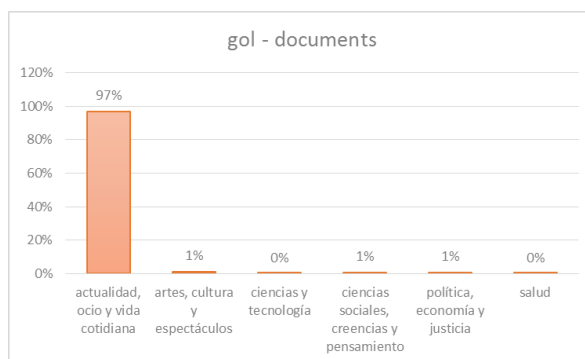
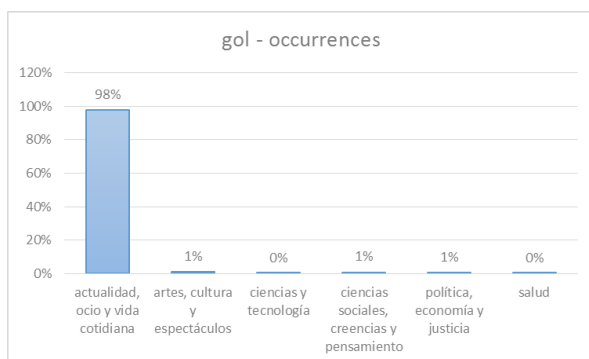


Absolute frequency: 5,115

N. of documents: 3,045

Normalised frequency: 20.68 hits per million words

3. Results

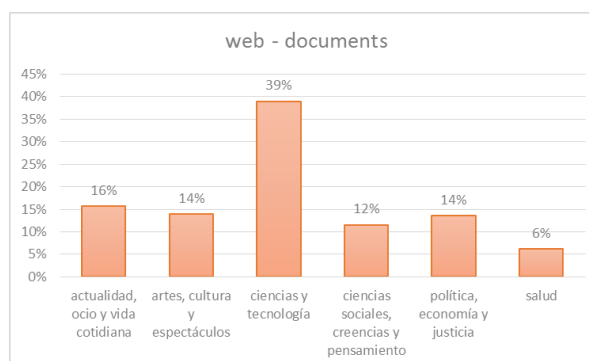
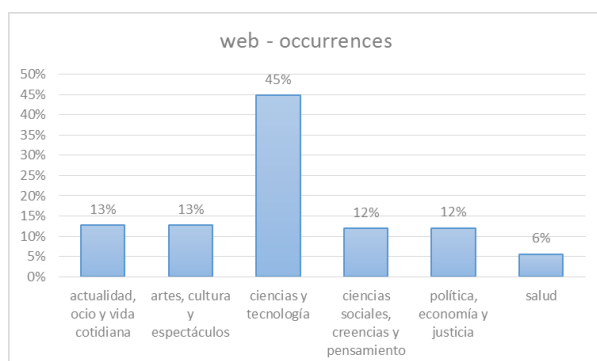


Absolute frequency: 4,300

N. of documents: 2,169

Normalised frequency: 17.38 hits per million words

3. Results

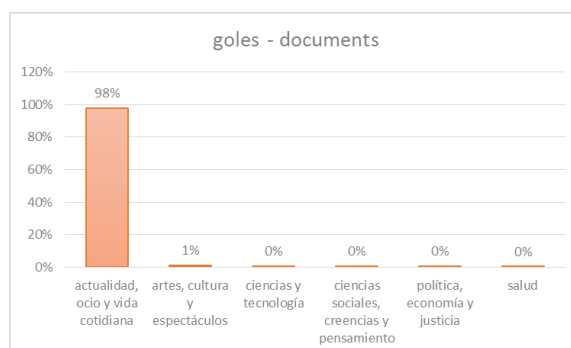
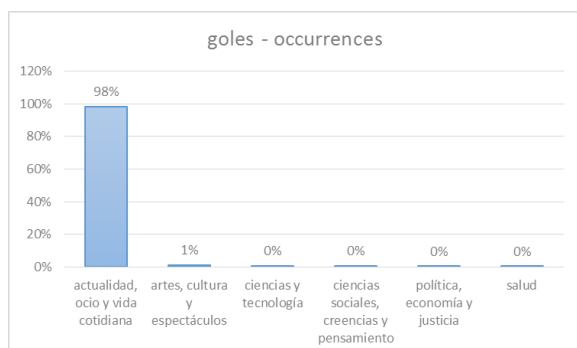


Absolute frequency: 4,302

N. of documents: 2,679

Normalised frequency: 17.39 hits per million words

3. Results

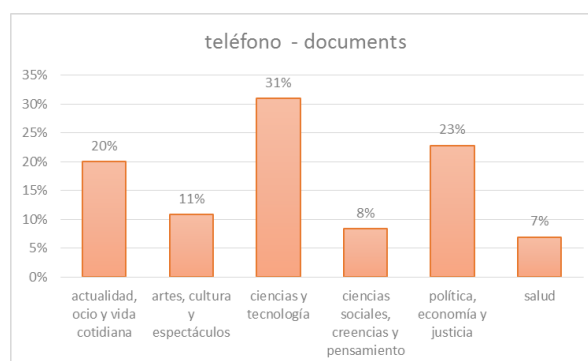
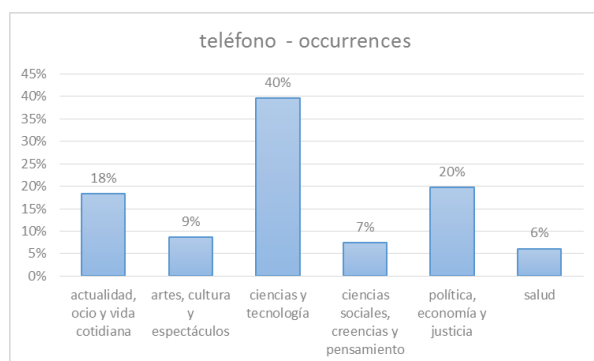


Absolute frequency: 2,983

N. of documents: 1,771

Normalised frequency: 12.06 hits per million words

3. Results

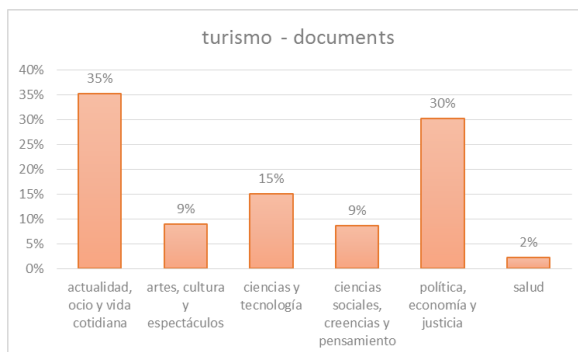
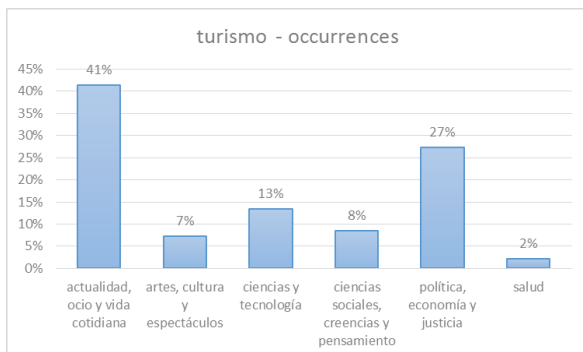


Absolute frequency: 2,787

N. of documents: 1,877

Normalised frequency: 11.27 hits per million words

3. Results

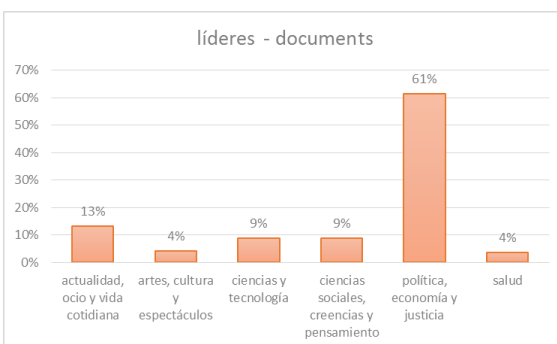
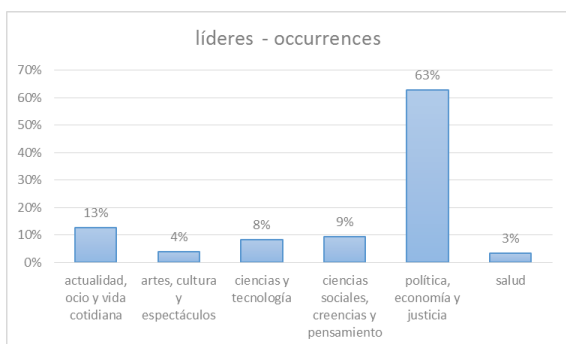


Absolute frequency: 2,787

N. of documents: 1,666

Normalised frequency: 11.27 hits per million words

3. Results



Absolute frequency: 1,665

N. of documents: 1,338

Normalised frequency: 6.73 hits per million words

4. Discussion

ANGLICISMS	MOST COMMON THEMATIC FIELD
club	Actualidad, ocio y vida cotidiana
fútbol	
gol	
goles	
turismo	
televisión	Artes, cultura y espectáculos
web	Ciencias y tecnología
teléfono	
-	Ciencias sociales, creencias y pensamiento
líder	Política, economía y justicia
líderes	
-	Salud

4. Discussion

- Not important differences “occurrences” – “documents”
- Difference between the most common thematic area and the rest of them: very big in words belonging to the sports field: fútbol, gol, goles, (club)
- Great difference between the usage of the singular form “líder” in the field “actualidad, ocio y vida cotidiana” (29%) and its plural, “líderes”, in the same thematic sphere (13%)
 - Analysis of specific concordances to find out possible reasons

5. Conclusion

- Online corpora: useful resource for the study of Anglicisms in a recipient language.
- CORPES XXI: insights into the usage of English loanwords in Spanish in the 21st century.
- Thematic distribution of the ten most common Anglicisms in the Spanish press.
- Future avenues: looking into the concordances of some words (líder – líderes)



contexts



patterns, possible reasons for relevant usage differences

6. References

- Andersen, G. (2005). Assessing algorithms for automatic extraction of Anglicisms in Norwegian texts. *Corpus Linguistics* 2005. Retrieved from <http://www.birmingham.ac.uk/Documents/college-artslaw/corpus/conferencearchives/2005-journal/ContrastiveCorpusLinguistics/Assessingalgorithms.doc>
- Andersen, G. (2011). Corpora as lexicographical basis: The case of anglicisms in Norwegian. *VARIENG - Studies in Variation, Contacts and Change in English, Methodological and Historical Dimensions of Corpus Linguistics* (6). Retrieved from <http://www.helsinki.fi/varieng/series/index.html>
- Andersen, G. (2012a). A corpus-based study of the adaptation of English import words in Norwegian. In *Exploring Newspaper Language: Using the Web to Create and Investigate a Large Corpus of Modern Norwegian* (pp. 157–192). Amsterdam / Philadelphia: John Benjamins Publishing Company. Retrieved from https://books.google.es/books?hl=es&lr=&id=RJmPfmQq_20C&oi=fnd&pg=PA157&dq=%22A+corpusbased+study+of+the+adaptation+of+English+import+words+in+Norwegian%22&ots=6lNI2gE4d3&sig=vzbCe__eWdfCrOY9okSbGZRBdOc
- Andersen, G. (2012b). Semi-automatic approaches to Anglicism detection in Norwegian corpus data. In *The Anglicization of European Lexis* (Cristiano Furiassi, Virginia Pulcini, Félix Rodríguez González (eds.), pp. 111–130). Amsterdam / Philadelphia: John Benjamins Publishing Company. Retrieved from https://books.google.es/books?hl=es&lr=&id=nI3y6TJw4YIC&oi=fnd&pg=PA111&ots=tlybNy6vyt&sig=r8jtJ9Iz5m6_jr2iJz8G9P-Yca
- Laviosa, S. (2007). Studying Anglicisms with Comparable and Parallel Corpora. *Belgian Journal of Linguistics*, 21 (The Study of Language and Translation. Vandeweghe, Willy, Vandepitte, Sonia, and Van de Velde, Marc (eds.)), 123–136.

6. References

- Núñez Nogueroles, E.E. (forthcoming), *Anglicisms in the Spanish Contemporary Press / El anglicismo en la prensa española contemporánea*, Granada: Universidad de Granada. Tesis Doctorales.
- Oncins Martínez, J. L. (2009a). Why to 'fill a niche' in Spanish is not a bad thing anymore: A corpus-based look at Anglicisms in Spanish. In *Proceedings of the Corpus Linguistics Conference 2009 (CL2009)*, 2009, pág. 396 (p. 396). Retrieved from <https://dialnet.unirioja.es/servlet/articulo?codigo=4012466>
- Oncins-Martínez, J. L. (2009b). Towards a Corpus-Based Analysis of Anglicisms in Spanish: A Case Study. *International Journal of English Studies*, 115–132.
- Oncins-Martínez, J. L. (2012). Newly-coined Anglicisms in contemporary Spanish. A corpus-based approach. In *The Anglicization of European Lexis* (pp. 217–238). Amsterdam / Philadelphia: John Benjamins Publishing Company. Retrieved from https://books.google.es/books?hl=es&lr=&id=nl3y6TJw4YIC&oi=fnd&pg=PA217&dq=%22Newlycoined+Anglicisms+in+contemporary+Spanish%22&ots=tlycSx2DA&sig=AyRzAzX7_U17kLG3hQrdVXlvR1E
- Seidel, U. (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). The University of Alabama TUSCALOOSA. Retrieved from http://libcontent1.lib.ua.edu/content/u0015/0000001/0000441/u0015_0000001_0000441.pdf

Thank you for your attention