

Alicante Journal of English Studies

Revista Alicantina de Estudios Ingleses

n° 30, December 2017



Departamento de Filología Inglesa
Universidad de Alicante

Alicante Journal of English Studies Revista Alicantina de Estudios Ingleses

ISSN 0214-4808 • CODEN RAEIEX

Issue Editors
Editors

José A. Sánchez and Ignacio M. Palacios
Isabel Balteiro and Miguel Ángel Campos

Editorial Board

Juan Carlos Acuña Fariña (University of Santiago de Compostela) • Asunción Alba (UNED) • Román Álvarez (University of Salamanca) • Bernd Dietz (University of La Laguna) • Fernando Galván (University of Alcalá) • Ramón López Ortega (University of Extremadura) • Susana Onega (University of Zaragoza) • Francisco Ruiz de Mendoza Ibáñez (University of La Rioja)

Advisory Board

Manuel Almagro Jiménez (University of Seville) • José Antonio Álvarez Amorós (University of Alicante) • Fernando Cerezal (University of Alcalá) • Ángeles de la Concha (UNED) • Helena Calsamiglia (Pompeu Fabra University) • Javier Díaz Noci (University of the Basque Country) • Teresa Gibert Maceda (UNED) • José S. Gómez Soliño (University of La Laguna) • Shaeda Isani (Université Grenoble-Alpes) • Sergio Maruenda Bataller (University of Valencia) • José Mateo Martínez (University of Alicante) • Ana Isabel Ojea López (University of Oviedo) • Ignacio Palacios Martínez (University of Santiago de Compostela) • Hanna Skorczynska Sznajder (Valencia Polytechnic University) • María Socorro Suárez Lafuente (University of Oviedo) • Justine Tally (University of La Laguna) • Jeroen Vandaele (University of Ghent) • M. Carmen África Vidal (University of Salamanca) • Francisco Yus Ramos (University of Alicante)

This volume has been funded by the Grants for the Publication of Scientific Journals from the Office of the Vice President of Research and Knowledge Transfer of the University of Alicante for the Promotion of R&D&I.

Contents

Acknowledgements 3

Introduction: Gaining insights into English as a contact language and its diffusion
Palacios Martínez, Ignacio M. and Sánchez Fajardo, José A. 5-22

ESSAYS

Part I Theoretical and empirical contributions to the study of English contact in World Englishes

English as a Lingua Franca (ELF) in ESP contexts. Students' attitudes towards non-native speech and analysis of teaching materials
Calvo-Benzies, Yolanda Joy 27-60

Rhoticity in Chinese English: An experimental investigation on the realization of the variant (r) in an Expanding Circle variety
Li, Zeyu and Kabak, Baris 61-91

Language contact in Gibraltar English: A pilot study with ICE-GBR
Loureiro-Porto, Lucía and Suárez-Gómez, Cristina 93-119

Contact-induced variation in clausal verb complementation: The case of *regret* in World Englishes
P. Romasanta, Raquel 121-147

Syntactic complexity and language contact: A corpus-based study of relative clauses in British English and Indian English
Tamaredo, Iván 149-182

Part II: Exploring trends and methodologies in the study of anglicized lexis

English lexical items in Egyptian Arabic. Code-switching or borrowing?
Kniaz, Malgorzata 185-210

How *VIPs* and *IT-professionals* behave in Russian. The use of loan abbreviations in Russian analytical composites
Tarasova, Elizaveta 211-236

Present-day Spanish fashion lexicon dresses up in English Rodríguez Arrizabalaga, Beatriz	239-276
Analysis of the presence of anglicisms in a Spanish internet forum: Some terms from the fields of fashion, beauty and leisure Luján García, Carmen	277-300
The influence of the English language on the description of cosmetic products Tejedor Martínez, Cristina	303-329
Exploring the dynamics of English/Spanish codeswitching in a written corpus González Cruz, María Isabel	331-355

Acknowledgments

We would like to express our sincere gratitude to the general editors of the *Revista Alicantina de Estudios Ingleses*, Dr. Isabel Balteiro and Dr. Miguel Angel Campos Pardillos, for inviting us to edit this special issue. Their help and guidance have been invaluable throughout the process.

Special thanks also to all the contributors, without whose work this issue would not have been possible. Thanks are also due to the various reviewers who, despite their heavy schedules and commitments, devoted their time to the reviewing process and provided very valuable comments and suggestions for improvement: Isabel Balteiro Fernández (U. of Alicante), Mario Cal Varela (U. of Santiago), Marisa Díez Arroyo (U. of Oviedo), Xabier Fernández Polo (U. of Santiago), Cristiano Furiassi (U. of Torino), Sebastian Knospe (U. of Greifswald), Antonio Lillo Buades (U. of Alicante), Lucía Loureiro Porto (U. of Balearic Islands), Paloma Núñez Pertejo (U. of Santiago), Javier Pérez Guerra (U. of Vigo), José Carlos Prado Alonso (U. of Oviedo), Virginia Pulcini (U. of Torino), Esperanza Rama Martínez (U. of Vigo), Félix Rodríguez González (U. of Alicante), Paula Rodríguez Abruñeiras (U. of Valencia), Paula Rodríguez Puente (U. of Oviedo), Elena Seoane Posse (U. of Vigo), Cristina Suárez Gómez (U. of Balearic Islands), and Nuria Yáñez Bouza (U. of Vigo).

We are also grateful to the Spanish Ministry of Economy and Competitiveness (grant FFI2015-64057) and the Galician Ministry of Innovation and Industry (grants CN/2015/004 and ED431D 2017/09).

Introduction: Gaining insights into English as a contact language and its diffusion

Ignacio M. Palacios Martínez
Universidade de Santiago de Compostela
ignacio.palacios@usc.es

José A. Sánchez Fajardo
Universidad de Alicante
jasanchez@ua.es

The aims of this volume are twofold: to contribute to the study of English as a contact language and its various manifestations in World Englishes, and to explore the causes and effects of the influence and diffusion of English in several languages, with particular reference to Spanish.

As Schreier and Hundt (2013: 1) have noted, the English language “has been contact derived from its very beginnings” and to this we can add that due to its rapid and far-reaching extension, leading to its current role as a global contact language (Görlach, 2002), it continues to be closely connected to a wide range of communities of speakers and languages across the world. In fact, as Onysko (2016: 192) claims, “the notion of language contact emerges as a valid candidate for being a unifying characteristic of all Englishes”.

The effects of this language contact can sometimes be seen in the influence of the substrate(s) language on new varieties of English, bringing about changes and innovations in the system. However, not all changes and innovations can be attributed exclusively to the role of the substrate element, since cognitive and transfer processes may also play a part here. Furthermore, sociolinguistic factors such as language attitudes and group identity might also be in operation. In broad terms, the outcomes of language contact can be seen in a number of contact-induced phenomena, including borrowing, phonological change, language transfer, relexification, code mixing, creole formation and code-switching, among others (Sankoff, 2001; Winford, 2005).

The contributions in this volume will thus address the following issues of English as a contact language and its diffusion: syntactic complexity and language contact in traditional and New Englishes, English as a lingua franca in ESP contexts, contact-induced variation in clausal verb complementation in World Englishes, factors determining the degree of rhoticity in Expanding Circle Englishes, language contact in Gibraltar English, the influence of English in Arabic word-formation, the presence of English on internet forums and in the description of fashion and cosmetic products



addressed to speakers of Spanish, the dynamics of English/Spanish code-switching and anglicisms in twenty-first century Spanish press.

In the light of this, we have organised the chapters into two main sections. The first will deal with theoretical and empirical issues in the study of English contact in World Englishes, and the second will reflect new trends and methodologies in the understanding of anglicized lexis, looking at particular genres and different languages where the presence of English is clearly observed.

1. Theoretical and empirical contributions to the study of English contact in World Englishes

A detailed account of the state of the art regarding the nature and evolution of World Englishes is beyond the scope of this introduction. (See, to mention just a few, Kirkpatrick, 2007; Kachru et al., 2009; Melchers and Shaw, 2011; Schneider, 2011, 2013; Buschfeld et al., 2014; Schreier, 2014; Seoane and Suárez-Gómez, 2016). Here we will simply mention some of the main issues as a means of introducing, contextualizing and better understanding the studies included in this section.

First of all, a note on terminology. We have already used several terms, World Englishes, New Englishes, Expanding Circle Englishes, to refer to what is essentially the same concept: those varieties of English that have emerged in different parts of the world, more particularly in Africa and Asia, and which in many cases have become national languages with their own status in the administration, political and educational domains of the respective countries. Other terms used on the same lines include Postcolonial Englishes, Extraterritorial Englishes, Indigenized Englishes, NIVES (Non-Native Institutionalized Varieties of English) and Nativized Varieties. These do not always have precisely the same meaning, with some carrying specific nuances, although on the whole they share the same core sense (See Schenider, 2013: 132-133, for a discussion of this issue). In this volume we will generally use World Englishes, in that it is considered to be one of the most neutral terms.

The field of World Englishes is now a well-established one, to the extent that several international journals exist (*English World Wide*, *World Englishes-Journal of English as an International and Intranational Language*), as well as a number of handbooks and collected volumes (see above for some of these), an academic organization, the International Association for World Englishes (IAWE), conferences, and a very considerable amount of internationally published scholarly work on different issues in the field.

Corpus linguistics has played a central role in the study of World Englishes and, in this respect, the ICE project (*International Corpus of English*) including twenty-six components on specific varieties (Canada, East Africa, Hong Kong, India, Ireland, Jamaica, New Zealand, the Philippines, etc.) has made it possible to conduct comparative studies, since all components share the same size, one million words, and the compilation follows a similar format in terms of corpus design and grammatical annotation. Indeed, several subcorpora of ICE have been used by contributors in this volume, such as

Tamaredo-Meira and Loureiro-Porto & Suárez-Gómez, and these latter authors are themselves presently compiling the ICE component on the variety of English spoken in Gibraltar.

Another of our contributors, P. Romasanta, uses the *Corpus of Global Web-based English* (GloWbe), created by Mark Davies and released in 2013. Containing some 1.9 billion words of blog texts from twenty different countries, it is a very useful tool for research, not least because of its size. The international organization of Corpus Linguistics ICAME (International Computer Archive of Modern and Medieval English)¹ also offers several corpora of this kind, such as the *Australian Corpus of English* (ACE), the *Wellington Corpus of Written and Spoken English* (New Zealand) and the *Kolhapur Corpus of Indian English*. To this list we can add a number of diachronic corpora, such as ARCHER (*A Representative Corpus of Historical English Registers*), the *Corpus of Early New Zealand English* (CENZE), the *Corpus of Early Nineteenth-Century Ontario Newspaper English*, the multigeneric *Strathy Corpus of Canadian English*, the *Old Bailey Corpus*, and the *Historical Corpus of Singapore English*, to mention just a few. (See Collins, 2015 for further information on all these corpora, plus findings from grammar studies that use them).

In addition to this, the *Electronic Atlas of the Varieties of English* (eWAVE 2.0, November 2013), designed and compiled at the University of Freiburg, Germany, by Kortmann and Lunkenheimer, is an extremely useful tool for research in this area, since it provides information on 235 morphosyntactic features in a total of 76 varieties of English, including traditional dialects, high-contact mother-tongue Englishes, indigenized second-language Englishes and English-based Pidgins and Creoles.²

We should also make brief reference to the main models that have been proposed to explain the development of World Englishes and to classify all these varieties (See Kirpatrick, 2007: 30-35, for more detailed information). Kachru's (1985) 'Three Circles' model is perhaps regarded as being the most traditional. This model establishes British English and American English within the first circle, Philippines English, South-African English and Indian English, for example, are classified in the second circle, whereas varieties of English associated with China, Russia or Korean would be included in the Expanding Circle. It has been quite controversial in that it is static and thus cannot easily account for the dynamism of many World Englishes, where the role and status of varieties may change in a relative short space of time. Secondly, as Kirpatrick claims (2007: 29-30), the model "underestimated the roles that English would come to play in Expanding Circle countries". A more comprehensive and more widely accepted approach is that of Schneider's (2003) dynamic model, in which five main phases or stages are identified: foundation, exonormative stabilisation, nativization, endonormative stabilization and differentiation. The first of these, foundation, refers to the phase in which English establishes itself in a country where this did not previously exist. Exonormative stabilization is said to be the stage in which the new variety follows the pattern introduced by the settlers. In the third phase, nativization, the two existing varieties come together and the English language is heavily restructured, particularly regarding vocabulary and grammar. This leads to the following stage, endonormative stabilization, in which the

new variety is gradually accepted by the community. Finally, in the differentiation period the new variety is widely used, reflecting the local community and culture. More recently, Mair (2016) has applied de Swaan's systems-theoretical model of global multilingualism to World Englishes to set out a new proposal, one which draws our attention to recent aspects of globalization illustrated by the migrations produced following the Cold War, the influence of the entertainment industry, and the revolution in communications resulting from the rise of the Internet. In this model Standard American English is considered to be "the hub of the World System of Englishes" (2016: 24).

This brief consideration and contextualization of the field leads us on to the contributions in the current volume. The first study of this section, by Yolanda Joy Calvo Benzies, of the University of Balearic Islands (pages 27-60), deals with student attitudes towards non-native speech in ESP teaching materials. Her approach combines pedagogical and attitudinal perspectives on non-native accents in the context of a new use and variety of the English language, English as a Lingua Franca (ELF), which itself can be seen as a contact language, in that it is adopted by different groups of speakers with their own native languages who decide to make use of a common code for communicative purposes.

Yolanda Calvo's paper is in fact divided into two main parts. In the first, she looks at the presence of native and non-native accents in ESP textbooks through an analysis of the CD tracks from six ESP textbooks for students of Law, Business and Tourism. Her results show, as expected, that native speakers prevail (70%) over non-natives ones, while less than 5% of the tracks contain only non-native speakers. On the whole, the combination of both native and non-native types of speakers seems to be the norm, and in broad terms non-native accents are quite well presented in these textbooks. From this, and bearing in mind the limitations of the sample considered, one might suppose that a change in the general approach to English language teaching is now gradually taking place, with teachers exposing their students to non-native accents and not limiting themselves to native speakers of English as the model to follow.

The second part of Calvo's study focusses on students' attitudes towards these non-native varieties. For this purpose, two questionnaires, adopting a Likert scale from 1 to 10, were administered to a total of 14 Law and Tourism ESP students. Five specific non-native accents were selected, these being from native speakers of German, French, Chinese, Polish and Spanish. Results indicate that all subjects regard having good pronunciation of English very positively and they all maintain that it is not possible to have a good command of spoken English if one does not know how to pronounce words well.

As regards attitudes to non-native accents, Tourism students seem to be more flexible to such accents while Law students value native accents more markedly. Considering the specific non-native accents, speakers of English, German and Polish are the most highly rated by these students while French and Spanish speakers receive the most negative responses. All in all, this study draws our attention to the relevance of including audio materials of non-native speakers of English in general and ESP courses. The study also shows that the emergence of a new variety of English such as ELF may have a direct

impact on the learning materials and, as a result, on the whole teaching process, particularly in regard to language activities and the standard to be used as a model.

The paper by Li and Kabak, of the Universities of Münster and Würzburg (Germany) respectively (pages 61-91), discusses the realization of postvocalic /r/ in Chinese English, often seen as one of the most notable phonotactic differences in accents of English. However, while rhotic sounds have been widely researched in the Inner and Outer Circle varieties, this is not the case in Expanding Circle Englishes, such as Chinese English. This study, then, is a very useful contribution in this area.

The authors start by describing in detail the variant /r/ in Inner Circle, Outer Circle and Expanding Circle Englishes, before focussing on rhoticity in Mandarin Chinese as a means of understanding its influence in Chinese English. The following section presents the methodology used, which is based on the elicitation of data from a total of 13 teachers of English by means of 3 instruments: a reading aloud task (a short story), a free speech task in English on one or two topics from a list provided, and reading a text aloud in Mandarin Chinese. This was complemented by a questionnaire in which participants were asked to provide personal information about their learning of English and other languages, and their preferences and attitudes towards a particular standard (British English versus American English).

A total of 762 instances of the production of /r/ were elicited, and these were then coded according to several demographic and speaker-specific variables (age, language attitude, L1 Mandarin rhoticity, speech style, additional language(s) spoken, etc.) and linguistic factor groups (preceding vowel, stress and syllable structure).

The results show that Chinese English, as represented by these 13 teachers of English, is only slightly rhotic. Considerable variation across participants is found, with a proportion of rhoticity ranging between 40% to 60%. A multivariate analysis of the data also reveals that the preceding vowel is the most significant influencing factor here, while stress and syllable structure were not significant. Moreover, those subjects with American English as target norm tended to produce rhoticity while this was not the case with those who opted for the British English standard. Speech style is also a determining factor, since Chinese English speakers produce more rhotic sounds in free speech than in the reading task. This, in the authors' views, could be related to the degree of formality of the two tasks. By contrast, rhoticity in the L1 does not seem to exert as significant an influence as expected.

The authors conclude by suggesting that in broad terms Chinese English is closer to the American English accent, and they relate their findings, in line with Mair's theory (see above), to the Dynamic System Theory (DST) approach to multilingualism and second language acquisition, in which languages are seen to be constantly evolving and developing. This leads them to formulate a new model for sound patterns, the 'Dynamic Equational Approach', in which sound patterns are characterized by a degree of gradience and variability as the result of the interplay of a number of language-internal and extralinguistic factors.

The paper by Lucía Loureiro-Porto and Cristina Suárez-Gómez (pages 93-119), both senior lecturers of English at the University of Balearic Islands, deals with a variety of

English that is very close to us, the English spoken in Gibraltar. As they explain, this variety of English is particularly interesting in that it has been in close contact for centuries not only with Spanish but also with Italian, Hebrew and Arabic.

The study is organised into two main parts. In the first of these, the authors describe the current state of the compilation of the Gibraltar component of the International Corpus of English (ICE) which they are compiling together with Elena Seoane, University of Vigo, and Jennifer Ballantine Perera, University of Gibraltar. This team was commissioned in 2014 to compile the ICE-GBR and thus far 100,000 words have been collected. It is also worth noting in this respect that the compilation of this corpus will fill an important gap in the literature and in World Englishes research in general, since there is currently no similar corpus that allows for the empirical study of this variety with hard, reliable and representative data.

In the second part of the study, several morphosyntactic features are analysed by looking at the extent to which contact with Spanish is indeed present and can be identified. For this purpose, they analyse the section on press new reports in the corpus, amounting to a total of 40,000 words, and compare it with a sample of a similar size and nature from the British English component of the ICE corpus (ICE-GB). Four main issues are examined: the frequency of the passive voice, relative markers, the use of titles (*Mr, Dr, Sir*) and pseudo-titles (*Tory, therapist*), and code-switching.

The results show that in the case of the passive voice the low frequency of this construction in Gibraltar English may be explained by contact with Spanish, in which it is not very common. Regarding relativizers, the authors hypothesize that the high frequency of the relative *que* in Spanish could favour the use of its counterpart in English *that*. The data analyzed reveal that this can be applied to those examples with inanimate antecedents but not to those with animate ones. However, the latter tendency has been identified quite extensively worldwide, and thus in this case contact with Spanish as an explanation should be taken with care. As regards the use of titles and pseudo-titles, it was assumed that due to the influence of Spanish a lower frequency of these items in ICE-GBR than in ICE-GB would be detected. However, the opposite is true and the data show a significantly higher proportion of members of this category in Gibraltar English. Finally, several examples of code-switching were recorded in the press material studied. In general these uses were used for expressive purposes and were restricted to a few nouns such as *alcalde* and constructions such as *mano a mano*. These cases of code-switching are regarded by the authors as the clearest examples of the influence of Spanish on the variety of English spoken in Gibraltar.

The paper by Raquel P. Romasanta, of the University of Vigo (pages 121-147), is concerned with clausal verbal complementation in World Englishes, taking the verb *regret* as a case in point. Language contact and the causes and effects of this language contact constitute the starting-point of the study, in that the author aims to explore the extent to which the preference for finite patterns over non-finite structures in World Englishes (British English, American English, Jamaican English, Hong Kong English and Nigerian English, in particular) is conditioned by transfer processes from the substrate languages or by cognitive mechanisms derived from the language contact situation itself.

The results, based on the analysis of 14,984 attestations of this verb recorded from the General and Blogs components of the GloWbE (Global Web-based English corpus),³ show that the six possible complementation patterns identified by FrameNet (a lexical database of English based on examples of how words are actually used)⁴ for the verb *regret* (NP, *wh*-clauses, *to*-infinitive, *-ing*, *that* and *zero*) are all present in the five varieties of English considered. The use of a prepositional phrase as a complement of *regret*, as in *regret about*, *regret for*, is also reported, a finding which is not mentioned in general reference grammars or on FrameNet. Moreover, while British English and American English show similar distributions regarding finite versus non-finite complementation patterns, this is not true for the non-native varieties, these differences being highly significant statistically. Jamaican English seems to be the closest in its distribution patterns to the native varieties, followed by Hong-Kong English and Nigerian English.

The data also reveal that while British English and American English show a clear preference for non-finite patterns, this is not the case for the non-native varieties. As regards the effects of language contact on this feature, the author concludes that the absence of non-finite patterns in Cantonese may be responsible for the low frequency of finite and non-finite constructions found in Hong-Kong English. Something similar can be said of Nigerian English, since none of the three languages spoken in this country (Hausa, Igbo and Yoruba), or even French, the most popular foreign language learnt in Nigerian schools, have gerund forms. The case of Jamaican English is slightly different since the number of non-finite forms here is quite high, although this variety finds itself in an advanced phase of evolution as compared to the other non-natives ones considered.

Together with the influence of the substrate, the role of cognitive processes is also detected, in the form of hyperclarity and isomorphism. Thus, several variables and factors can be said to exist in a state of interplay in the phenomenon of clause contact, with sociolinguistic and cognitive probably two of the most significant.

Iván Tamaredo, research scholar at the University of Santiago de Compostela, focuses in his paper (pages 149-182) on the study of relative clauses in British English and Indian English from the perspective of syntactic complexity. For this, he makes use of a corpus-based methodology analysing data extracted from the British and Indian components of the International Corpus of English. British English here represents a native variety while Indian English corresponds to a L2 variety which emerged under contact conditions and with a high exonormative influence.

The author defines the concept and the metrics of syntactic complexity by referring to previous descriptions in the literature, then turns to the complexity of relative clauses in particular. It is widely acknowledged that relative clauses are difficult to process, since they involve a relation of dependency between a gap and an antecedent, to the extent that it is possible to speak of an 'Accessibility Hierarchy', which means that languages differ with respect to the way they relativize.

After a general description of relative clauses in the two varieties selected, Tamaredo presents the main results of the study, derived from the analysis of over one thousand instances in 40 written and spoken texts from each of the ICE components. A total of 7

variables are considered (variety, text type, degree of restrictiveness, relativizer, relativizer position, preposition placement and relative clause complexity) and the data obtained are analyzed by means of a series of hierarchical configurational frequency operations which reveal that British English shows a preference for covert relativizers (*that* and *zero*) in restrictive relative clauses, with a difference between speech and writing observed: while the former favours *that*, the latter prefers *zero*. In contrast, in Indian English overt relativizers, that is, *wh*-pronouns, are the preferred option. The results here include two findings which are considered unexpected: the high proportion of *zero* relatives in written informal texts in the two varieties studied, and the abundance of *that* relatives in spoken formal texts in Indian English.

The paper concludes by claiming that differences found with respect to relative formation in British English and Indian English can be accounted for by differences in syntactic complexity: speakers of Indian English are more prone to produce simpler relative clauses than British English speakers and they also tend to shun complex relatives in prepositional complement and genitive positions. We can say, then, that their condition as L2/L3 speakers of English may also explain their preference for relative clauses, which may be easier to process and understand.

2. Exploring trends and methodologies in the study of anglicized lexis

The rise of English as a *lingua franca* has been associated with the influx of anglicized lexis and phrases. Whether a ‘plagued’ word stock (cf. Graddol, 2006), or just a result of over-lexicalization, these English-induced loanwords have come into existence mainly due to an ingrained sense of terminological ‘gap-filling’ (*pen-drive*, *chat*), unknown referents in the target language (cultural borrowing), or merely a sociolinguistic restriction (slangy or jargoned language). However, despite these reasons, it is true to say that the vast majority of languages (if not all those which are connected to a globalized market) have been influenced by English at some point.

The diffusion of English has been traditionally linked to empowering Anglophone territories, and to their political and economic influence on neighboring countries. The United Kingdom and the United States, since the peak of the Industrial Revolution in the 19th century, have laid the groundwork for an ensuing standardization of technical and trade-related terms. As expected, technological advances and globalization have given way to an extensive importation of ‘the foreign’ and the assimilation of ‘the anglicized native’. Also, the dissemination of pop culture and mass media have greatly contributed to an era that has been described as ‘English-dominated’. No doubt, the enactment of English as a communicative means in academic programs such as CLIL (cf. Marsh, 2002; Coyle et al., 2010) has guaranteed its ever-increasing linguistic permeability, particularly in economic and political blocks such as the European Union (Görlach, 2002; Furiassi et al., 2012). With all this in mind, the impact of English on languages and cultures offers an unparalleled opportunity to study the variability of contact languages in more detail.

The processes of code-switching and linguistic borrowing in these languages in contact are a reflection of how the *nativization* of English terms and phrases might trigger

universal categories and trends. This is precisely where anglicists have focused their attention: the description and contrast of morpho-semantic variations in the target language due to the influence of English structures and patterns. This involves the implementation of specific methodological and typological procedures in the collection and analysis of data.

One of the most notable attempts to explore the process of linguistic borrowing is the *Loanword Typology Project* (henceforth *LWT*) (cf. Haspelmath and Tadmor, 2009), based on contributors from 41 languages. This non-English-based cross-linguistic approach, reflected in other, related works (Wichmann and Wohlgemuth, 2008; Wohlgemuth, 2009), has sought to describe the historical circumstances and evolution of loanwords. To put it another way, based on a fixed list of 1,460 semantic items (not words), contributors were given the task of indicating the existence of loanwords in their own languages. These counterparts were extremely useful as a means of identifying the historical evolution of words, and a variability scope of borrowing was established: 0- no evidence of borrowing, 1-very little evidence for borrowing, 3-perhaps borrowed, 4-probably borrowed, 5-clearly borrowed (Haspelmath and Tadmor, 2009: 13).

Thus, one of the breakthroughs of *LWT* involves unraveling the question of word-class dependency and lexical accommodation in the process of linguistic borrowing, by comparing the lexical systems of various recipient languages. The objective of this study is “to get a clearer idea of lexical borrowability by examining the loanwords in a reasonably representative and reasonably large set of languages (say, 30-40 languages), and by making inductive generalization over the data assembled in this way” (Haspelmath, 2008: 3). The onomasiological perspective of this novel methodology is reflected in the semantic arrangement of lexical items. By establishing an initial list of meanings (not words), comparable lexical samples could be gathered, which are processed in terms of semantic word class, typical context, and morpho-syntactic traits.

The definition of a loanword, however, reflects the typological restriction of the concept of linguistic borrowing: a loanword is defined “as a lexeme that has been transferred from one lect into another and it is used as a word (rather than as an affix, for example) in the recipient language” (Haspelmath and Tadmor, 2009: 13). Although this limitation guarantees a parallel quantitative study and a comprehensive visualization of language contact, it does not add structural and conceptual distinctions between the processes of calquing and loaning: “loanword (or lexical borrowing) is here defined as a word that at some point in the history of a language entered its lexicon as a result of borrowing (or transfer, or copying)” (Haspelmath, 2009: 36).

The second notable procedure worth mentioning here is corpus-based, and is intended to explore the linguistic impact of a source language (English) on a recipient one, provided that all types of lexical borrowing are collected and processed. Notwithstanding their limited scope (only two languages are involved), dictionary and corpus-based studies provide a general examination of language contact, including direct (loanword) and indirect borrowing (calque):

Through corpus-based research it is possible to carry out qualitative studies and improve microstructural features of general dictionaries and dictionaries of anglicisms, i.e. obtain

information about frequency, period of adoption, usage context and authentic examples. Corpus-driven methods facilitate quantitative studies and contribute to the improvement of the macrostructure of general dictionaries and dictionaries of anglicisms by detecting new candidate entries (Furiassi et al., 2012: 18).

Alternatively, the *GLAD (Global Anglicism Database) Network*⁵ is aimed at studying the influence of a common source language (English) on a number of recipient languages (Spanish, French, German, Danish, Chinese, etc.). Based on national corpora and a common typological strategy (Furiassi et al., 2012), *GLAD* seeks to set up a global database of English-induced lexical units, revealing how English has affected these languages at the lexical, morphological, syntactic and pragmatic levels. The results are intended to depict common grounds and dissimilarities in the Anglicization of lexis, particularly in terms of morpho-phonological and semantic variations. In spite of its lexicon-centered perspective, such multicultural corpora are useful tools in the examination of global trends and converging paradigms. In addition, a semasiological study of loanwords can shed more light on both the typology of intralinguistic word-formation mechanisms and the extrapolation of English patterns in the target language.

One of the most visible, and intriguing, results of this importation is perhaps the pseudo-Anglicization process. A pseudo-anglicism, also known as a ‘false’ loan, has traditionally been defined as “a word or an idiom that is recognizably English in its form (spelling, pronunciation, morphology, or at least one of three) but is accepted as an item in the vocabulary of the receptor language even though it does not exist or is used with a conspicuously different meaning in English” (Furiassi, 2010: 34). Other authors (Filipović, 1985; Onysko, 2007; Balteiro and Campos, 2012; Furiassi and Gottlieb, 2015) have also contributed to the study of these units from phonological, morphological and semantic perspectives, particularly in the examination of European languages.

The richness of false loans lies in their inner expression of lexical creativity and interlinguistic transmutation, i.e. the acquisition of source-language traits and their combination with target-language ones. Also, pseudo-anglicization is characterized by the unlikelihood of English speakers detecting the meaning of the existing words, which might itself constitute a useful and delimiting way of looking at the phenomenon from the point of view of the ‘monoglot’ Anglophone reader or listener (cf. Furiassi and Gottlieb, 2015): *footing* ‘jogging’ in Spanish, *Handy* ‘mobile phone’ in German, *caddy* ‘type of trousers’ in Polish, to name a few.

Like anglicisms in general, the adoption of false loans is motivated by a number of factors: language economy, in which brevity is key to understanding the coinage of shortened forms (cf. Furiassi and Gottlieb, 2015), and the “taste for the exotic, the charm of a foreign language, and the glamorous quirk of being creative and playing with language are the core motivations for the birth of false Anglicisms” (Furiassi, 2010: 62-63). Thus, the complexity of loanwords depends on word choice, semantic cues and sociolect. In other words, referential or speaker-related motivations can lead to the formation of ‘new’ words in the target language, with some updated underlying pragmatic and semantic features.

As discussed above, Spanish and English have long been regarded as languages in ever-increasing contact. The reasons are once again associated with geographical location, economic influence and immigration. In the nineteenth century, as Britain emerged as a world power, English began to exert a great influence on Spanish as it did on other European languages, and this influence increased greatly leading to the replacement of French as the main source of foreign loans in Spanish (cf. Rodríguez, 2002: 128). But it was not until the twentieth century that Spanish underwent a major Anglicization of its lexis, particularly due to the influence of an empowered United States, which wielded its considerable emerging economic and political influence over its neighboring countries on the American continents.

Technology is one of the domains that has characterized this predominantly southbound influence. Its adoption was induced by the rapid importation of referents and concepts. The ephemerality and haste in adopting this technolect by Spanish speakers has led to semantic extension and variation: *bisnes* (<*business*) is used colloquially for ‘exchange of goods’, but is also restricted to ‘illegal business’ in some Spanish-speaking South American countries. Such a transition denotes the semantic obscurity by which slang or jargon loanword stock is characterized, and the relation between pragmatic factors and diachronic shift.

The influence that English has had on Spanish is undeniable, reflected especially in the degree of ‘cultural borrowing’ (Weinreich, 1953; Gómez Capuz, 1998; Hoffer, 2002; Treffers-Daller, 2010). First defined as “loaned objects and practices that pass from one community to the other (cultural diffusion)” (Bloomfield, 1933: 444), a cultural loan is not another type of linguistic borrowing, but rather a transversal concept that depends on the features of the imported referent. The existence or not of referents or concepts in the target language determines the nature of a cultural loan. *Halloween* and *laptop* are unadapted anglicisms used in European Spanish, but they differ in the type of referent denoted: whereas *Halloween* is imported to convey a non-existent concept in Spanish, *laptop* coexists with *portátil* and is restricted to a specific technolect or jargon. The study of the pragmlinguistic dimensions or categories of cultural borrowing (cf. Sánchez, forthcoming 2018) is a reliable means of measuring the index of cultural density, that is, the impact that certain languages might have on their neighboring ones.

Simultaneously, a vast number of Spanish words or *hispanicisms* have been imported into English. Approximately a thousand of these loanwords have been attested in the *Oxford English Dictionary* (Rodríguez, 2001: 84), and this number is yet higher in American English, unlike in other English variants. Hispanic or Latino migration plays a fundamental role in this borrowing process, which is particularly tangible in bordering territories of North-America such as New Mexico, Florida and California. These territories abound in Spanish-origin toponyms: *Los Angeles*, *San Francisco*, *Santa Barbara*, etc., which have left an imprint of over two hundred years of Spanish dominance in the area.

The import of Spanish words, as above, is frequently conditioned by the lack of referents in English, particularly when referring to non-American cultural traits: *sombrero* ‘a Mexican hat’, *abogado* ‘lawyer specialized in immigration cases’, *taco*,

tamale (<*tamal*). The impact of the Hispanicization of English is also visible in the coinage of colloquial phrases, in which these Spanish or hybrid units are motivated by a sense of ‘intentional disguise’ (Galinsky, 1964) or marginalization: *zero, nothing, nada; mi casa essu casa*.

However, it is the emergence of *Spanglish* what has intrigued Spanish and English sociolinguists most, due to the creation of common lexical and syntactic patterns. Regardless of its acceptance or so-called intrusiveness, Spanglish has generally been recognized as “a pidgin, or creole language; an interlanguage; or an Anglicized Spanish language” (Neuliep, 2015: 105). The nature of Spanglish does not rely to any great extent on the intricate and diverse forms of code-switching and lexical borrowing, but on the communicative and pragmatic functions conveyed by what has been called “a dynamic fusion of crashing cultures noticeably merging at the interface of language” (Rothman and Rell, 2005: 516). Following Jakobson’s functional framework, Appel and Muysken (cf. 2005) have associated the six language functions with the switching process:

- (a) *referential*: the non-existence of concepts and items in the target language, or the sense of confidence of bilingual users when using certain words in specific communicative settings.
- (b) *directive*: the exclusion or inclusion of hearers in a conversation.
- (c) *expressive*: speech characterized by stretches of mixed language, in which two identities are easily recognized.
- (d) *phatic*: changes in the tone of a conversation, these in accordance with the language used.
- (e) *metalinguistic*: allusions to both languages directly or indirectly, depending on the message.
- (f) *poetic*: the use of stylistic tropes in both languages, these intended to make use of bilingual puns, jokes, etc.

The study of the contact between Spanish and English is of great relevance due to the sociolinguistic and historical repercussions underlying the processes of code-switching and borrowing. Two of its more practical contributions here are the examination of hispanicisms and anglicisms in both languages, and the analysis of how these imported units are restricted variably to certain lects. Spanglish is perhaps the most palpable interlinguistic (Spanish and English) construct that results from bilingualism and acculturation. Its functions and trends might shed more light on the understanding of analogous mixtures such as “Taglish (Tagalog-English in the Philippines), Hinglish (Hindi-English in India), Franglais (mixture of French and English), Portuñol/ Portunhol (Portuguese-Spanish), Guarañol (Guaraní-Spanish), and many others” (Lipski, 2008:40).

Given the comprehensive dimension of the contact of English and other languages, this volume on ‘World Englishes’ and the diffusion of English, as discussed above, intends to go beyond the examination of existing varieties of English, and to present readers with some analysis of the Anglicization of languages, in particular Spanish. As such, this section (Part II) of the special issue of *Revista Alicantina de Estudios Ingleses*

(RAEI) presents studies on the influence of English on dissimilar languages such as Spanish, Arabic and Russian. These articles contribute to a better understanding of specific linguistic cases and their methodological principles, towards an understanding of what type of converging foreign trends are imported, and more importantly, how these structures and patterns are adapted in the target or recipient language. Much attention has been devoted here to Spanish, with four studies that show how these variations occur in different semantic fields, such as fashion or the media, and how the process of establishing a database could itself be driven by the features of the word-stock in question.

With theoretical issues in mind, this part of the volume opens with the paper “English lexical items in Egyptian Arabic. Code-switching or borrowing?” (pages 185-210), in which Malgorzata Kniaż, of the Jagiellonian University, explores the conceptual differences of code-switches and borrowings in an analysis of 3,443 bilingual projections of complementizers. The use of the Matrix Language Model reveals how disparities in the results are a reflection of the situation of English in the Arab world. The article includes an in-depth study on the phenomena of code-switching and borrowing, which are seen explicitly as a continuum, as suggested by previous work in the field (cf. Toribio, 2015).

An outstanding aspect of this article is the correlation between the structural characteristics of languages in contact and the sociolinguistic contexts of speakers. More specifically, in this case study it is found that those speakers lacking knowledge of Standard Arabic are more prone to fill lexical gaps through the acquisition of English loanwords, and the shortage of morphological integration is driven by the visible incongruence of these two languages, as opposed to other Indo-European ones, such as French, German, Spanish, etc.

Following this, Elizaveta Tarasova, of the Institute of the Pacific United in New Zealand, focuses on the assimilation of morphological patterns from English into Russian, particularly denominal compound nouns. Based on Krysin’s principle (1975), the author asks whether these loanwords are still productive and how instances of English morphology might be duplicated. Interestingly, Tarasova introduces her study by providing non-Russian speakers with a thorough explanation of how N + N compounding occurs in Russian morphology, which clarifies some analogous word-formation mechanisms in the two languages.

In terms of morphology, the author makes a comprehensive analysis of the nature of compound bases and their derivational paradigms in the target language. The Russification of imported units is also examined as a mechanism of semantic variation, in that some of the resulting constructs seem to be “stylistically coloured and belong to expressive vocabulary” (pages 211-236). Her study not only seeks to confirm whether N + N is still a productive pattern in Russian, but also provides readers with a valuable analytical procedure that might be used for the analysis of other borrowed composites.

As suggested above, this volume is especially aimed at presenting current research on the Anglicization of Spanish lexis. To have a better understanding of how different analytical procedures can lead to the examination of same-domain loanwords, two articles on fashion and beauty have been included: “Present-day Spanish Fashion Lexicon

Dresses up in English”, by Beatriz Rodríguez Arrizabalaga”, of the University of Huelva (Spain), and “Analysis of the Presence of Anglicisms in an Internet Forum: Some Terms of the Fields of Fashion, Beauty, and Leisure”, by Carmen Luján García, of the University of Las Palmas de Gran Canarias (Spain).

In the examination of Rodríguez’s findings (pages 239-276), readers are given a valuable analysis of the impact of English on the media and advertising. This is a necessary step in understanding the motivations underlying the diffusion of an anglicized lexicon in the Spanish fashion industry. Although the paper centers on a corpus-based analysis of 14 anglicized lemmas related to the domain of ‘fashion’, its relevance lies in the study of these ‘superfluous’ loanwords (cf. Pratt, 1980; Lorenzo, 1996) and their autochthonous Spanish equivalents. The resulting quantitative data shows a diachronic progression of lexical frequency and word variation, illustrating the ever-increasing dominance of gratuitous anglicisms in the fields of marketing and sales.

Similarly, Luján’s paper (pages 277-300) looks at the semantic fields of ‘fashion’, ‘beauty’ and ‘leisure’ through the analysis of corpus-based data. The study is divided into three overall stages: (i) extraction of English-induced lemmas, (ii) classification of anglicisms (Furiassi and Gottlieb, 2015), and (iii) exploration of their morpho-syntactic peculiarities. The final stage provides valuable reflections on the semantic change undergone by some of the lemmas that have been classed as ‘pseudo loan’, and the type of orthographical and phonological variations that might be involved. The use of quantitative data and authentic contexts are of great value here in supporting some of the findings with regard to borrowing typology and frequency.

Following these studies on fashion-related loanwords from diachronic and synchronic perspectives, Cristina Tejedor Martínez, of the University of Alcalá (Spain), also uses a corpus-based procedure to describe the process of Anglicization in the domain of cosmetics. Her paper, “Mixing English and Borrowings in the Description of Cosmetics Products” (pages 303-329), is based on the compilation of English-induced loanwords, previously retrieved from authentic texts used in the advertising campaigns of well-known cosmetic brands in Spain. The resulting corpus is examined and contrasted by using well-known prescriptive dictionaries of both English and Spanish, which allows for a clear-cut distinction of anglicisms and pseudo loans. One early quantitative finding indicates that cosmetology loans are not as frequent as other productive domains in Spanish, such as computing and tourism, although a significant number of them are indeed found. A morphological exploration reveals a duplicity of English word-formation patterns, which is complemented by some comments on the psychological impact of these words on consumers.

The corpus-based studies in this section, though apparently similar in the scope of the research, are effective in the analysis of anglicized lexis from various perspectives: semantic equivalency, morpho-syntactic variation, dia-synchronic evolution, etc. Not only do their findings add clarity to our understanding of linguistic borrowing in Spanish, but they also complement a multilinguistic examination of the target language with brief pragmatic commentaries on speakers’ intentions and the nature of referents.

As noted above, the presence of Spanish-induced words in English is also a result of the process of language contact, although it has traditionally been related to historical and social phenomena, such as Spanish settlements in North America and migration. M^a Isabel González Cruz, of the University of Las Palmas de Gran Canarias (Spain), discusses the presence and diffusion of Hispanicisms in two samples of romances, as part of a more ambitious project devoted to the compilation of a corpus of these lexical borrowings.

The examination of 16 romances is aimed at exploring the socio-pragmatic cues and discourse functions that characterize the processes of code-switching and lexical borrowing. By studying the implicatures of borrowing in literary texts, the author examines the predominant type of code-switching according to Lipski's classification model (1982). González Cruz's study, "Exploring the dynamics of English/Spanish code-switching in a written corpus" (pages 331-355), uses literary sources because they afford a clear perspective on the variability of code-switches in the target language. In doing so, the study contributes to a practical exploration of how these underlying motivations are associated with the degree of borrowability and the typology of the units collected.

Language contact and the diffusion of English are necessarily intertwined. The anglicizing impact of this *lingua franca* results in multifaceted variation in the target languages, and in a pragmatic and semantic transmutation of native units. As such, it represents a complex area of research. Exploring analytical procedures and analogous word-formation patterns contributes to a greater understanding of both the gradable influence of languages in contact and the pragmatic values of anglicized lexis and syntax in a globalized world. We commend all the authors in this volume for their valuable contributions to the field.

Notes

1. See: <<http://clu.uni.no/icame/>>
2. See: <<http://ewave-atlas.org/>> for further information.
3. See: <<https://corpus.byu.edu/glowbe/>>.
4. <<https://framenet.icsi.berkeley.edu/fndrupal/>>
5. <<https://www.nhh.no/en/research-centres/global-anglicism-database-network/>>

References

- Appel, Rene and Pieter Muysken (2005): *Language Contact and Bilingualism*. Amsterdam: Amsterdam University Press.
- Balteiro, Isabel and Miguel A. Campos (2012): "False anglicisms in the Spanish language of fashion and beauty". *Ibérica*, 24: 233-260.
- Bloomfield, Leonard (1933): *Language*. New York: Holt, Rinehart and Winston.
- Buchsfeld, Sarah, Thomas Hoffmann, Magnus Huber & Alexander Kautzsch, eds., (2014): *The Evolution of Englishes: The Dynamic Model and Beyond* [Varieties of English around the

- World G49]. Amsterdam & Philadelphia: John Benjamins.
- Collins, Peter, ed. (2015): *Grammatical Change in English World-Wide*. Amsterdam & Philadelphia: John Benjamins.
- Coyle, Do, Phillip Hood and David Marsh (2010): *CLIL*. Cambridge: Cambridge University Press.
- Furiassi, Cristiano (2010): *False anglicisms in Italian*. Milano: Polimetrica.
- Furiassi, Cristiano, Virginia Pulcini and Félix Rodríguez (2012): *The Anglicization of European Lexis*. Amsterdam & Philadelphia: John Benjamins.
- Furiassi, Cristiano and Henrik Gottlieb, eds. (2015): *Pseudo-English. Studies on False Anglicisms in Europe*. Berlin: De Gruyter Mouton.
- Galinsky, Hans (1964): "Stylistic aspects of linguistic borrowing. A stylistic and comparative view of American elements in Modern German and British English". In H. G. Lunt, ed., *Proceedings of the 9th International Congress of Linguistics*. Cambridge, Mass., August 27-31, 1962. The Hague: Mouton, 374-381.
- Gómez Capuz, Juan (1998): *El préstamo lingüístico. Conceptos, problemas y métodos*. Valencia: Universidad de Valencia.
- Görlach, Manfred (2002): *English in Europe*. Oxford: Oxford University Press.
- Graddol, David (2006): *English Next: Why Global English May Mean 'the End of English as a Foreign Language'*. London: British Council.
- Filipović, Rudolf (1985): "Pseudoanglicisms in European languages". In U. Pieper and G. Stickel, eds., *Studia Linguistica Diachronica et Synchronica. Werner Winter Sexagenario*. New York & Berlin: De Gruyter, 249-255.
- Haspelmath, Martin (2008): "Loanword typology: Steps toward a systematic cross-linguistic study of lexical borrowability". In T. Stolz, D. Bakker and R. Salas Palomo, eds., *Aspects of Language Contact: New Theoretical, Methodological and Empirical Findings with Special Focus on Romancisation Processes*. Berlin: Mouton de Gruyter, 43-62.
- Haspelmath, Martin (2009): "Lexical borrowing: concepts and issues". In M. Haspelmath and Uri Tadmor, eds., *Loanwords in the World's Languages: A Comparative Handbook*, 35-54. Berlin: De Gruyter Mouton.
- Haspelmath, Martin and Uri Tadmor (2009): "The loanword typology project and the world loanword database". In M. Haspelmath and U. Tadmor, eds., *Loanwords in the World's Languages: A Comparative Handbook*. Berlin: De Gruyter Mouton, 1-34.
- Hoffer, Bates L. (2002): "Language borrowing and language diffusion: An overview". *Intercultural Communication Studies*, 11(4): 1-37.
- Kachru, Braj B. 1985: "Standards, codification and sociolinguistic realism: The English language in the Outer Circle". In R. Quirk and H. Widdowson, eds. *English in the World*. Cambridge: Cambridge University Press, 11-30.
- Kachru, Braj B., Yamuna Kachru and Cecil L. Nelson, eds. (2009): *The Handbook of World Englishes*. Oxford: Blackwell.
- Kirkpatrick, Andy (2007): *World Englishes: Implications for International Communication and English Language Teaching*. Cambridge: Cambridge University Press.
- Krysin, Leonid (1975): "Stupeni morfemnoi chlenimosti inoyazychnyh slov". In E. Zemskaya, ed. *Razvitie sovremennogo russkogo yazyka. Slovoobrazovanie. Chlenimost' slova*. Moscow: Prosveshchenie, 227-231.
- Lipski, John M. (1982): "Spanish English language switching in speech and literature – theories and models". *The Bilingual Review*, 9(3): 191-212.
- Lipski, John M. (2008): *Varieties of Spanish in the United States*. Washington, D.C.: Georgetown

- University Press.
- Lorenzo, Emilio (1996): *Anglicismos Hispánicos*. Madrid: Gredos.
- Mair, Christian (2016): “Beyond and between the “Three Circles”: World English research in the age of globalisation”. In E. Seoane and C. Suárez. *World Englishes. New Theoretical and Methodological Considerations*. Amsterdam & Philadelphia: John Benjamins, 17-35.
- Marsh, David, ed. (2002): *CLIL/EMILE- The European Dimension: Actions, Trends and Foresight Potential*. Brussels: European Commission.
- Melchers, Gunnel and Philip Shaw (2011): *World Englishes* [The English Language Series]. London: Hodder Education.
- Neuliep, James W. (2015): *Intercultural Communication. A Contextual Approach*. Thousand Oaks: Sage Publications.
- Onysko, Alexander (2007): *Anglicisms in German: Borrowing, Lexical Productivity and Written Codeswitching*. Berlin & New York: De Gruyter.
- Onysko, Alexander (2016): “Language contact and world Englishes”. *World Englishes. Language Contact and World Englishes*. Special Issue 35(2): 191-195.
- Pratt, Chris (1980): *El Anglicismo en el español peninsular contemporáneo*. Madrid: Gredos.
- Rodríguez González, Félix (2001): “Spanish contribution to American English word-stock: An overview”. *Atlantis*, 23(2): 83-90.
- Rodríguez González, Félix (2002): “Spanish”. In M. Görlach, ed. *English in Europe*. Oxford: Oxford University Press, 128-150.
- Rothman, Jason and Amy B. Rell (2005): “A linguistic analysis of Spanglish. Relating language to identity”. *Linguistics and Human Sciences*, 1(3): 515-536.
- Sánchez Fajardo, José A. (Forthcoming 2018): “The pragmalinguistic dimensions of cultural borrowing”. *International Journal of Language and Culture*. Amsterdam & Philadelphia: John Benjamins.
- Sankoff, Gilliam (2001). “Linguistic outcomes of language contact”. In P. Trudgill, J. Chambers and N. Schilling-Estes, eds., *Handbook of Sociolinguistics*. Oxford: Basil Blackwell, 638-688.
- Schneider, Edgar W. (2003): “The dynamics of New Englishes: From identity construction to dialect rebirth”. *Language*, 79(2): 233-281.
- Schneider, Edgar W. (2011): *English Around the World: An Introduction*. Cambridge: Cambridge University Press.
- Schneider, Edgar W. (2013): “English as a contact language”. In D. Schreier and M. Hundt, eds. *English as a Contact Language*. Cambridge: Cambridge University Press, 131-148.
- Schreier, Daniel (2014): *Variation and Change in English: An Introduction*. Berlin: Eric Schmidt.
- Schreier, Daniel and Marianne Hundt, eds. (2013): *English as a Contact Language*. Cambridge: Cambridge University Press.
- Seoane, Elena and Cristina Suárez-Gómez, eds. (2016): *World Englishes. New Theoretical and Methodological Considerations*. Amsterdam & Philadelphia: John Benjamins.
- Toribio, Almeida J. (2015): “Code-switching among US Latinos”. In M. Díaz-Campos, ed. *The Handbook of Hispanic Sociolinguistics*. Malden (Mass.): Wiley Blackwell, 530-552.
- Treffers-Daller, Jeanine (2010): “Borrowing”. In M. Fried, J.O. Östman and J. Verschueren, eds. *Handbook of Pragmatics*. Amsterdam & Philadelphia: John Benjamins, 17-35.
- Weinreich, Uriel (1953): *Languages in Contact*. The Hague: Mouton.
- Wichmann, Søren and Jan Wohlgemuth (2008): “Loan verbs in a typological perspective”. In T. Stolz, D. Bakker and R. Salas Palomo, eds., *Aspects of Language Contact. New Theoretical, Methodological and Empirical Findings with Special Focus on Romancisation Processes*. Berlin & New York: Mouton de Gruyter, 89-121.

Winford, Donald (2005): "Contact-induced changes: classification and processes". *Diachronica*, 22: 373-427.

Wohlgemuth, Jan (2009): *A Typology of Verbal Borrowings*. Berlin & New York: Mouton de Gruyter.

Essays

Part I

Theoretical and empirical contributions to
the study of English contact in World
Englishes

English as a Lingua Franca (ELF) in ESP contexts. Students' attitudes towards non-native speech and analysis of teaching materials¹

Yolanda Joy Calvo-Benzies
Universidad de las Islas Baleares
yolandajoy.calvo@uib.es

ABSTRACT

This paper focuses on non-native accents in ESP classrooms. In particular it looks at native and non-native speakers of English accents used in the audio material accompanying six ESP textbooks. In a second study, a group of undergraduate ESP students of Law and Tourism were asked to assess some of the non-native speakers accents found in these materials, focussing on aspects such as fluency, pronunciation, intelligibility and foreign accent. More specifically, they were asked to rate the following non-native accents of speakers in English: French, German, Polish, Chinese and Spanish. Results from the first part of the study show that native speaker models continue to be present in ESP textbooks to a far higher degree than non-native ones. In the second part, the non-native accents that students rated most positively were those of German and Polish speakers, and those seen in the most negative terms were French and Spanish. In general, the Law students tended to value native accents more than non-native ones, whereas students of Tourism broadly accept both native and non-native accents.

Keywords: ELF, ESP, textbook analysis, students' attitudes, native and non-native accents

1. Introduction and statement of purpose

There is no doubt that English has been for some time -and still is- the main language used in many “different domains around the globe” (Bayyurt and Sifakis, 2017: 4) such as in international situations of commerce, business or politics. As Dogancay and Hardman (2017: 19) point out, this “global spread of English has led to the emergence of diverse varieties of English that represent different sociocultural norms, political affiliations and bilingual/multilingual identities”. In other words, this expansion has had a profound effect on the world we live in; in fact, a whole new context of English has recently emerged, one in which a broad range of variability in the language is possible, sometimes known *English as a Lingua Franca* (Seidlhofer, 2001; Jenkins, 2007; Walker, 2010; Seidlhofer, 2011; Mauranen, 2012; Motschenbacher, 2013; Mauranen, et al., 2015; Mauranen et al., 2016).

This new context in the use of the English language, which will be referred to as ELF from now onwards, has greatly influenced the field of language learning and teaching, especially regarding the speaking and listening skills as well as pronunciation. Concerning the latter, in recent decades, there have been many changes regarding: a) the amount of attention paid to this language area, and, b) the methods used to teach it (see Celce et al. (2010) for detailed descriptions on the wide variety of teaching methods available to teach pronunciation). Following Grant (2014), we can nowadays distinguish two main approaches to the teaching of this area of language. On the one hand, as summarised in Table 1 below, *Traditional Approaches* can be characterised as having as their main aim a perfect or native-like pronunciation, with teachers being exclusively native speakers of the target language, and with the speaker models used in the classroom again being of a native origin. By contrast, *Current Approaches*, also summarised in Table 1, emphasise what Grant (2014) calls *comfortable intelligibility*. Both native and non-native speakers are considered appropriate teachers in this area, as long as they are proficient speakers of the target language; in addition, a wide variety of models and standards can be distinguished here, according to issues such as context, the listener or the ultimate purpose of instruction.

	Traditional approaches	Current approaches
Learner goals	Perfect, native-like pronunciation	Comfortable intelligibility
Language background of teachers	Native-speaking teachers	Native-speaking and proficient non-native speaking teachers
Speaking models	Native-speaker models	Variety of models and standards depending on the listener, context, and purpose

Table 1: Traditional and current approaches to pronunciation teaching (adapted from Grant, 2014: 6)

On these general lines, Walker (2010: 5) describes how “the goal in pronunciation teaching has been for learners to achieve a native-speaker accent” and “the two dominant models used to this end have been the standard British accent, RP (Received Pronunciation), and the standard US accent, GA (General American)” (Walker, 2010: 5); whereas nowadays, although speakers who use ELF to communicate with others “still want to be comfortably intelligible”, the people who determine whether they are intelligible or not “are fellow non-native speakers” (Walker, 2014: 8). Another direct effect of the development of contexts like ELF is the fact that nowadays

English language teachers, both native and nonnative speakers of the language, need to know about varieties of English that they and their students are likely to encounter in and outside of classrooms, and they need to teach their students the sociolinguistics tools to navigate across Englishes” (Dogancay and Hardman, 2017: 19).

In terms of the present paper, a crucial issue in the teaching of pronunciation is the shift from the aim of sounding native-like when speaking a foreign language, to the idea that one can communicate with a foreign accent in the foreign language as long as the speaker is understood by listeners. In other words, the notion has arisen that speaking English with a non-native accent is considered to be acceptable, provided that communication is not inhibited, that is, that the speaker is wholly intelligible (Walker, 2010).

As a consequence of all of these changes, it is not surprising that a great deal of research over recent decades has focused on the presence of non-native accents within the classroom; in this area, many studies have looked at student attitudes to native and/or

non-native accents. Such studies can be divided into two main groups: a) studies on students' beliefs about native and non-native teachers, such as Lasagabaster and Sierra (2002), Madrid and Pérez-Cañado (2004), Cheung (2010), Alseweed (2012), Chun (2014), Walkinshaw and Hoang (2014), Dweik and Al-Barghouthi (2014), Buckingham (2015) and Karakas et al. (2016), to mention just a few; and, b) studies in which students listen to different native and/or non-native speakers and decide which accents they prefer. Within this latter group are studies by Jaber and Hussein (2011), Moinzadeh et al. (2012), Matura and Chiba (2014) and Thamer et al. (2014).

The findings from different studies analysing students' views on native and non-native teachers can be summarised as follows:

- a) Lasagabaster and Sierra (2002), Cheung (2010), Alseweed (2012), Dweik and Al-Barghouthi (2014), Chun (2014) and Walkinshaw and Hoang (2014) found that students believe both types of teachers have advantages. In general terms, native speakers are rated by students as being the best models for teaching spoken skills, pronunciation and cultural aspects; non-native ones, on the other hand, are believed to be better at teaching grammar. Moreover, the results reported in several of these studies indicate that students believe non-native speakers have more sympathy for them, since they are also non-native speakers and may resort to the learners' native language to resolve problems or misunderstandings with the foreign language. In other words, the non-native teachers were judged as "more effective in helping students with psychological aspects of language learning" (Chun, 2014).
- b) Lasagabaster and Sierra (2002), Madrid and Pérez-Cañado (2004), Alseweed (2012) and Karakas et al. (2016) found that the higher the level of proficiency of the students (especially those at university), the higher their preference for native English-speaking teachers.

As regards studies on students' attitudes to accents in English, Jaber and Hussein (2011) looked at the perceptions of native English speakers towards French, Japanese and Jordanian accented English, finding that "the Jordanian accent was considered as the most intelligible, followed by the French then the Japanese English accent" (Jaber and Hussein, 2011: 77). In Matura and Chiba (2014), standard Japanese speakers and speakers of accented Japanese were asked to assess several Outer Circle accents (Kenya, Ghana, Bengali, Hindi), that is, countries in which English is learned as a second language. Their findings indicated that the group with the most favourable reaction towards these Outer Circle accents were the speakers of accented Japanese. From the former, then, it seems that native speakers of English understand some non-native accents better than others. In contrast, in the latter study, the participants who reacted more positively towards the accent of second-language speakers of English were the non-native speakers of Japanese.

Similarly, two recently published studies focus on speakers' opinions to a number of native and non-native accents. Moinzadeh et al. (2012) compared the views of a group of Iranian students whilst listening to a lecture; some of these students listened to the lecture when it was given by a native American speaker, while the rest listened to the same

lecture read out aloud by a Persian speaker. The two groups then answered listening comprehension questions about the lecture. In turn, Thamer et al. (2014) took Malay, Malaysian-Chinese and Malaysian-Indian university students, and asked them to assess the accents of six speakers, all of whom were university lecturers. Three of these speakers were native speakers of the students' languages (Malay, Malaysian-Chinese and Malaysian-Indian), plus one Arabic speaker and two native speakers of English, one from Britain and the other from North America. In both of these studies, the speakers rated most positively by students were the non-native ones.

Another issue which is normally object of analysis when there are changes within the language learning and teaching field is material-assessment, as a means of gauging whether these are adequate for the approach followed, both inside and outside the classroom, or, on the other hand, whether changes here also need to be made. Despite the importance of this type of study, very little research seems to have been conducted on the presence of native and non-native speakers in the audio materials that typically accompany written textbooks. This lack of research may seem quite surprising since, as Matsuda (2017: xiv) explains, "one of the specific pedagogic recommendations" that approaches like ELF suggest is "increasing exposure to and raising awareness of diverse forms and functions of English" inside the classroom. One exception of the aforementioned lack of research on this topic is Kopperoinen (2011), who analysed several upper-secondary school EFL textbooks used in Finland, finding that non-native speakers were present in only 3% of the total listening materials in one group of textbooks assessed, and in just 1% in another group.

All in all, then, although some research has been conducted on native and non-native English speakers' attitudes towards other native and/or non-native speakers' accents, gaps remain regarding certain issues, such as the analysis of teaching materials. The present study aims to contribute here in three ways. First, whereas most existing research on native and non-native accents has focused on either English as a Foreign Language (EFL) or English as a Second Language (ESL), in the current paper the context will be English for Specific Purposes (ESP). Second, in studies looking at student preferences regarding native and non-native accents, the tendency has been to analyse the opinions of: a) native speakers after listening to non-native speakers (as in Jaber and Hussein (2011); or, b) non-native speakers assessing the accents of different native speakers (Matsuda and Chiba (2014), and also half of the subjects in Moizadeh et. al (2012). In the present paper, by contrast, Spanish non-native speakers of English give their opinions on the accents of other non-native speakers when speaking English. It is believed that conducting this type of study in a country like Spain may be interesting for two main reasons: a) thousands of tourists with many different L1s and/or L2s visit different parts of Spain almost all year round. For this reason, many Spaniards, especially those who work within the so-called *Sector Servicios*, end up using English to communicate with tourists from countries like Germany, France, China or Turkey, that is, both the foreigners and the Spaniards in these contexts interact by using a foreign language for them (in this case, English); b) there are also many Spaniards who travel abroad each year to countries like Germany or Poland; here, they also use English on many occasions to interact with

the locals since they do not know how to speak German or Polish, for example. Hence, in broad terms, English is used on many occasions for Spanish speakers to communicate with other non-native speakers of English. Finally, as mentioned above, very little work has so far been done on assessing textbooks to see the extent to which non-native accents are included in the accompanying audio materials. For this reason, part of the present paper will include an analysis on this precise issue: the role of native and non-native accents in some textbooks.

With these aims in mind, two empirical studies are described in this paper. More particularly, the first is concerned with the analysis of the audio(visual) materials which accompany textbooks used to teach ESP in different university disciplines (Business, Tourism and Law), in order to find out the extent to which speakers included in these materials are native or non-native. The sample selected represents materials recently being used with Law and Tourism undergraduate students. To complement this analysis, a group of undergraduate Law and Tourism ESP students at the University of the Balearic Islands were then asked to rate the speech of several non-native speakers extracted from these materials on aspects like fluency, pronunciation and intelligibility, as well as more subjective criteria such as *I wouldn't mind having an English teacher with this accent* or *this speaker sounds like a native speaker of English*. This second study was included because textbooks continue to be the main teaching materials used in language classes (Marks, 2006; Lopez-Jiménez, 2009); furthermore, students are clearly important participants in the classroom (Baker, 2011), and hence their points of view on the learning process should always be taken into account, including their perspectives on the teaching materials used therein.

2. Study 1: Presence of native and non-native accents in ESP textbooks

2.1. Research materials

The CD tracks from a total of six ESP textbooks were selected and analysed. As shown in Table 2 below, the sample is representative of currently available materials from major publishers (Cambridge University Press, Oxford University Press, Delta). In terms of the disciplines they focus on, textbooks 1-3 are addressed to Law students, textbooks 4 and 5 are mainly used to teach Business English (although, as will be explained below, can also be used to teach English to undergraduate Tourism students), and number 6 is intended exclusively for students of Tourism.

Number	Textbooks analysed	ESP discipline they focus on
1	English for Legal Professionals (2009, OUP)	Law
2	Absolute Legal English (2010, Delta Publishing)	Law
3	International Legal English CD1 + CD2 (2008, CUP)	Law
4	English for Meetings (2007, OUP)	Business English – Tourism
5	English for Job Hunting CD1 + CD2 (2008, OUP)	Business English – Tourism
6	Tourism 3 CD 1 + CD2 (2009, OUP)	Tourism

Table 2: List of ESP textbooks analysed

Although most of these materials can now be considered slightly out of date, given that they were published between 7 and 10 years ago, they were chosen because the majority of them are still in use on different ESP courses taught at the University of the Balearic Islands. Whilst some of them are the main course book used in a specific ESP subject, sections and activities from others are extracted to complement the main teaching materials. Thus, the third textbook here, *Tourism 3*, is the main book used in the third-year compulsory subject *English III* within the degree of Tourism; however, since this subject revolves around job interviews and business meetings, some sections from textbooks 4 and 5 are also used. In future studies, it may be interesting to analyse more-recently published materials which perhaps follow a more up-to-date approach to the teaching of ESP with non-native Englishes; in other words, the results obtained in the analysis of this sample of textbooks should not be considered as generalisations beyond this local context but rather as a preliminary analysis of the materials currently being used to teach ESP within the degrees of Tourism and Law at the University of the Balearic Islands.

The textbooks selected for analysis were divided into two groups: a) Law textbooks (Group A); and b) Tourism and Business textbooks (Group B). These two groups will be used for both empirical studies reported in this paper.

2.2. Data analysis

Once the materials had been selected, a database was created to enter the information extracted from the tracks found on the CD(s) accompanying each textbook. An example of the database designed is given in Table 3, below. The track number and page reference for each listening activity in the course books were recorded in separate columns. In the next column the number of speakers featuring in each audio excerpt was noted, these classified according to whether they were: a) native or b) non-native speakers of English. Finally, a field was also included for more detailed comments on a particular track, for

instance, whether information is given either in the textbook or on the track itself concerning the origins of a particular non-native speaker. More detailed tables were subsequently made as a means of comparing the tracks from each textbook (see section 2.3).

Track number	Page number	Accents of speakers		Comments
		Native	Non-native	
2	7	1	1	Japanese
3	8	1	1	Japanese
4	19	1	1	French
5	24	2	0	
6	31	7	0	

Table 3: Analysis of audio tracks from *English for Legal Professionals*

It is important to note here that the same speaker sometimes appeared on more than one track. However, since each of the tracks were considered separately, such speakers were counted as new each time. In other words, if the voice of the same German man featured on Tracks 4 and 16 in a particular textbook, Track 4 was analysed as containing a non-native speaker, and Track 16 as another example of a file containing a non-native speaker.

In principle all the audio tracks on an accompanying CD or CDs under analysis were considered. Nevertheless, in the case of Textbook number 6, *Tourism 3*, 18 tracks were excluded from the final analysis since they represented activities taken from the *pronunciation section* in the textbook; because these tracks were based on speakers simply reading aloud English words containing certain sounds, they were not considered as instances of running speech to be analysed here. The total number of tracks analysed was hence 213. Table 4 below sets out the number of tracks on the CD(s) for each textbook (column 4), the number of tracks finally analysed from each textbook (column 5), and the total number of tracks analysed in the two groups: a) Law textbooks; and, b) Business and Tourism textbooks (column 6). As can be seen in Table 4, the number of tracks analysed for each group of textbooks is quite homogeneous: 106 for Group A (Law textbooks) and 107 in Group B (Business and Tourism textbooks).

	Number	Textbook	Total number of tracks for each textbook		Number of tracks analysed		
GROUP A	1	English for Legal Professionals	14		14		Total number of tracks in law textbooks: 106
	2	Absolute Legal English	37		37		
	3	International Legal English	CD1 31	CD2 24	CD1 31	CD2 24	
GROUP B	4	English for Meetings	26		26		Number of tracks in Business and Tourism textbooks: 107
	5	English for Job Hunting	CD1 29	CD2 24	CD1 29	CD2 24	
	6	Tourism 3	CD1 23	CD2 23	CD1 15	CD2 13	

Total number of tracks analysed: 213

Table 4: Number of tracks on CD(s) of each textbook, plus total audio files analysed per group

2.3. Results

This section is divided into several parts. First, the general results for the analysis of all 213 tracks will be presented. Comparisons and differences between Group A and Group B will be made, that is, between Law textbooks versus Business-Tourism ones. Finally, the two textbook groups will be considered separately.

2.3.1. General results

As can be seen in Table 5 below, for the vast majority of the tracks analysed only native speakers of English were used (nearly 70%). Furthermore, both native and non-native speakers interacted together in a total of 57 tracks (over 26%). Finally, less than 5% of files extracted from the six ESP textbooks contained only non-native speakers. So, overall we see that non-native speakers are heard in only 30.9% of the tracks analysed.

In general terms, this first finding is perhaps quite surprising. As was noted in the Introduction, although non-native accents are broadly accepted today as correct accents of English (as long as the person communicates in an intelligible way), native speakers of English continue to be used as models in ESP textbooks much more frequently than non-native ones.

	Tracks with only native speakers	Tracks with only non-native speakers	Tracks with both native and non-native speakers
Total	147 (69.01%)	9 (4.23%)	57 (26.76%)

Table 5: Total number of tracks with only native speakers, with only non-native speakers, or with both

2.3.2. Law versus Business and Tourism textbooks

A comparison of the presence of native and non-native speakers in the audio files from the two groups of textbooks reveals several differences. First, as can be seen in Table 6 below, over 80% of the speakers featuring in tracks from Group A (Law) are native speakers of English, a figure which falls to 55% for Group B (Business and Tourism). In bare numbers, whereas non-native speakers feature in only 19 tracks (from a total 106) in Group A texts, this rises to 45 tracks in the case of Group B.

As will be discussed below when each of the textbooks is considered separately, the non-native speakers present in both groups of course books were either from different European countries (France, Germany or Poland, for instance) or from Asian countries (mainly China or Japan). It is also important to note that, although most of the speakers found on the tracks analysed were native speakers of English, these included people with a variety of different native accents. Thus, whereas some have a standard RP or GA accent, other speakers can be found with Australian English, Irish English, Scottish English or South African accents.

	Tracks with only native speakers	Tracks with only non-native speakers	Tracks with both native and non-native speakers
Group A textbooks	87 (82.08%)	7 (6.6%)	12 (11.32%)
Group B textbooks	60 (56.07%)	2 (1.87%)	45 (42.06%)

Table 6: Total number of tracks with only native speakers, only non-native speakers, or with both (Groups A and B)

2.3.3. Results obtained per discipline

2.3.3.1. Law textbooks

Table 7 sets out the use of native and non-native speakers in the three textbooks for Law students. As might be expected, given that over 80% of the tracks in the whole group here feature only native speakers, three of the four CDs in this group include a very high

percentage of tracks in which only native speakers of English appear,² over 80% in the Textbook *Absolute Legal English* and over 90% in the two CDs accompanying *International Legal English*.

The CD accompanying the textbook *English for Legal Professionals* contains quite a few examples of tracks in which both native and non-native speakers intervene together. The three non-native accents mainly represented here are Japanese, Spanish and French. By contrast, only seven of the total 37 tracks which appear on the CD for *Absolute Legal English* include non-native speakers, and these tend to be Polish and French speakers. Finally, the textbook *International Legal English* combines native and non-native speakers least often. As seen in Table 7, CD1 of this textbook does not feature a single example of a non-native speaker of English in its 31 tracks; similarly, 90% of tracks on CD2 feature only native speakers. One positive feature of this textbook, however, is that Track 15 of CD2 involves a South African native speaker of English, a native variety of English which is not present in most of the books analysed in this study.

Overall, we can say that the Law ESP books analysed do not give students the opportunity to listen to many speakers with non-native accents of English, something that goes against the current models of teaching English, that is, exposing students to people with different accents (Matsuda, 2017).

	Number of tracks analysed per textbook	Tracks with only native speakers	Tracks with only non-native speakers	Tracks with both native and non-native speakers
English for Legal Professionals	14	4 (28.57%)	0 (0%)	10 (71.49%)
Absolute Legal English	37	30 (81.08%)	6 (16.21%)	1 (2.7%)
International Legal English CD1	31	31 (100%)	0 (0%)	0 (0%)
International Legal English CD2	24	22 (91.66%)	1 (4.16%)	1 (4.16%)

Table 7: Complete analysis of Law textbooks

2.3.3.2. Business and Tourism textbooks

As with the previous group, two out of the three Business-Tourism textbooks analysed contain more examples of tracks in which only native speakers intervene than ones with non-native speakers. However, in one of these two textbooks, *Tourism 3*, the percentage of tracks with only native speakers is under 70% in both CDs, whilst it was always over 80% in the majority of the Law textbooks.

Non-native speakers do not appear in any of the 26 tracks with the book *English for Meetings*. By contrast, the book which best combines the presence of native and non-native speakers is *English for Job Hunting*, with non-native speakers appearing in over 55% of the tracks on CD1 and close to 80% in CD2. This textbook contains examples of speakers with a variety of different non-native accents, including people from Poland, Spain, China, Germany, Italy and The Netherlands.

Finally, CD1 of the textbook *Tourism 3* includes quite a few tracks with only native speakers of English, plus one featuring only non-native speakers. The situation in CD2 is somewhat different, because in nearly 70% of the tracks only native speakers of English appear, although there are four tracks featuring people with non-native accents. A disadvantage of this book is that little information is given on the origins of the non-native speakers, although some French and Spanish speakers are quite easy to identify. Also of note is the presence in CD2 of a native English speaker from South Africa, a variety which, as mentioned above for the textbook *International Legal English CD2*, is very rarely found in these course books.

Thus, it seems that, although Tourism and Business students once again tend to be exposed to native speaker accents more frequently than to non-native ones, some of the ESP textbooks for these disciplines do include quite a few examples of non-native accents.

	Number of tracks analysed per textbook	Tracks with only native speakers	Tracks with only non-native speakers	Tracks with both native and non-native speakers
English for Meetings	26	26 (100%)	0 (0%)	0 (0%)
English for Job Hunting CD1	29	12 (41.4%)	0 (0%)	17 (58.62%)
English for Job Hunting CD2	24	5 (20.83%)	0 (0%)	19 (79.16%)
Tourism 3 CD1	15	8 (53.33%)	1 (6.6%)	6 (40%)
Tourism 3 CD2	13	9 (69.23%)	1 (7.69%)	3 (23.08%)

Table 8: Complete analysis of the Business and Tourism textbooks.

3. Study 2: ESP students' attitudes towards non-native accents

3.1. Subjects

The participants in this study were Spanish ESP students. They were all bilingual speakers of Castilian and Catalan (Majorcan variety). Their age ranged from 20 to 23, with the exception of two participants who were over 40 years of age when the study was conducted. The students were drawn from two areas: Tourism and Law. All of these participants had studied English in their Primary, Secondary and Post-Secondary Education studies; hence, they had been exposed to EFL for approximately 10-12 years. Moreover, as will be mentioned below, they were all studying a course on ESP during the academic year 2016/2017. It was the first time the Law students were studying ESP whereas the Tourism ones had already studied ESP in the 2 previous academic years.

In an initial plan to assess Law students' views towards different non-native accents, volunteers were sought from students in the fourth-year optional subject, *English for Legal Purposes* (end of May 2017). Fewer than 10 participants in fact completed the tasks, so a group of former third-year Tourism students were invited to take part. In the end, as seen in Table 9 below, a total of 14 ESP students took part in the study, eight undergraduates of Law and six students currently enrolled in Tourism studies, all from the University of the Balearic Islands. Although the number of Law students who took part is slightly higher, two of these in fact only answered some general questions on native and non-native accents and did not assess any of the non-native speakers. Hence, a balanced group of 12 students completed all of the tasks, six from Law and six from Tourism.

Law students	Tourism students	Total
8	6	14

Table 9: Number of Law and Tourism students who took part in the study

We should note here that although the two groups of students (Law and Tourism) were studying English for particular purposes, the amount of exposure they had to English was quite different. For Law students at the University of the Balearic Islands, there is only one optional subject to study English. This course is offered to fourth-year students, and its aim is that they will achieve a B2 level in English. This is so that they will fulfil the obligatory requisite of having a B2 level in a foreign language when graduating in any university degree (according to the Bologna system). Undergraduates enrolled in the degree of Tourism and Hospitality, by contrast, have an ESP subject in every year of their degree: from first to third-year (*English I, English II, English III*) and in fourth year they can take an optional ESP subject (*English IV*).

3.2. Research materials: design and administration

This study seeks to explore ESP student views regarding non-native speakers of English. As a means of establishing connections between this and the previous study here, I used

some of the tracks described in the above analysis in which non-native speakers had been identified.

Extracts from all the audio material accompanying the textbooks in the previous study were selected, with the exception of *English for Meetings*, since in this case all of the speakers on the CD were native speakers of English.

The participants studying Law were asked to assess the accents of some of the non-native people identified in Group A textbooks, that is, those addressed to Law students, while the Tourism students were given tracks taken from the Business English and Tourism books. This approach was to ensure a high degree of general understanding of the tracks, in that Law students would be far more accustomed to listening to speakers discussing legal topics than those related to Business English or Tourism, and vice versa.³

Fortunately, the groups of textbooks analysed in the previous study often contained speakers with the same non-native accents. For example, Spanish, German and French speakers were found in both textbooks addressed to Law students and in the Tourism and Business materials. This allowed me to choose examples of the same non-native accents for both groups of students to assess; more specifically, the five non-native accents selected were: a) German, b) French, c) Chinese, d) Polish, and, e) Spanish (see Tables 10 and 11 below). Tracks in which the speaker did not give any information regarding his or her origins were chosen whenever possible, since one of the questions the students were asked when assessing each non-native speaker was *Where do you think this person is from?* The specific tracks each group of students were asked to assess can be found in Tables 10 and 11 below.

TRACKS FOR LAW STUDENTS TO ASSESS			
Speaker number	Book	Track number	Speaker's accent
1	Absolute Legal English	19	Polish man
2	Absolute Legal English	26	French man
3	English for Legal Professionals	2	Chinese woman
4	English for Legal Professionals	11	Spanish woman
5	International Legal English CD2	13	German man

Table 10: Non-native speakers assessed by Law students

TRACKS FOR LAW STUDENTS TO ASSESS			
Speaker number	Book	Track number	Speaker's accent
1	English for Job Hunting CD1	6	Polish man
2	English for Job Hunting CD1	8	German man
3	English for Job Hunting CD2	8	Chinese woman
4	Tourism 3 CD1	15	French woman
5	Tourism 3 CD1	18	Spanish man

Table 11: Non-native speakers assessed by Tourism students

Different questionnaires were designed for the tasks. To begin with, students were asked to fill out a questionnaire with some general questions on English pronunciation and on native and non-native accents. This followed a close-ended format in which the majority of the questions were presented using a Likert scale from 1 to 10, with 1 representing *I totally disagree* and 10 *I totally agree*. Moreover, in one question participants were asked to choose the native variety of English they believe they best understand.

The 10 questions in this questionnaire can be divided into two groups: most of the items revolve around students' opinions on general issues, such as the extent to which they believe English pronunciation is important, whether they think native-like English pronunciation is important or whether they believe that non-native speakers of English are able to speak and pronounce well, and so on. The second group of questions focus on students' preferences, namely, whether they prefer native teachers over non-native ones in their ESP classes, if they would rather listen to native speakers than to non-native ones in listening to audio material in the classroom, and the native accent they best understand. For reasons of space, the whole lists of questions can be found in Appendix 1.

In addition to this questionnaire, which dealt with basic or general issues, a further, 8-item questionnaire was created so that students could assess each of the non-native speakers on the recordings. The majority of these questions followed the same Likert scale design. Here students were asked to assess each speaker according to criteria such as *This speaker has a strong foreign accent when speaking in English, this person speaks English well, and I understand this person, they are intelligible*. This set of questionnaires also ended with an open question, in which students were asked to guess where each speaker was from (see Appendix 1 for the whole list of questions).

The questionnaires were administered online using Google Docs. A separate questionnaire was created for each of the 5 speakers to be assessed, so that participants did not have to carry out the whole series of assessments in one sitting. Each of these 5 questionnaires (one per non-native speaker), of course, contained exactly the same questions. Also, important to note is that all the questions were written in Spanish, one of all students' native languages, thus minimising the potential for ambiguity or misunderstandings.

3.3. Data analysis

As already described, participants were asked to fill in the same questionnaire on *General Aspects* and then a separate assessment of each of the five non-native speakers. However, two Law students failed to assess Speaker number 5, and two Law students only answered the *General Aspects* questionnaire, without assessing any of the speakers' accents. Due to the low number of fully completed questionnaires, I decided to include the answers of students who had only filled out some parts. Thus, in the results section, the number of responses will be slightly different in some sections than in others, as can be seen in Table 12.

	Number of Law students who filled out the questionnaire on...	Number of Tourism students who filled out the questionnaire on...	Total
General aspects	8	6	14
Speaker number 1	6	6	12
Speaker number 2	6	6	12
Speaker number 3	6	6	12
Speaker number 4	6	6	12
Speaker number 5	4	6	10

Table 12: Number of Law and Tourism students who answered each questionnaire.

As in the previous study, tables reflect students' answers in the different questionnaires. For reasons of space, the opinions given by each participant in each question cannot be discussed; therefore, in what follows the mean score for the students who actually responded in each case is given.

3.4. Results and discussion

In this section I will first discuss ESP students' general opinions on native and non-native accents, according to the answers given in the *General Aspects* questionnaire. As in the previous study, any differences of opinion observed across groups (Law and Tourism) will also be described. Participants' assessments of the non-native speakers will then be explored. As mentioned above, a total of 10 speakers were chosen, but each participant only had to assess five of them, these five depending on whether they were students of Law or Tourism. I will report students' views on each of the non-native accents: a) Polish, b) German, c) French, d) Chinese and e) Spanish. Then comparisons will be made across the five non-native accents to identify the one(s) said to be most and least intelligible, and the ones which, according to respondents, had the strongest or weakest foreign accent.

Before continuing, it is important to mention here that, although the sample is quite limited, the Chi-square test was used to see whether there were significant differences between subjects at the .0001 level. This statistical analysis was applied to the data

collected in Tables 13, 14, 18, 20 and 22. Nevertheless, as can be gathered from the tables in Appendix 2, none of these differences were statistically significant. Thus, detailed explanations on statistical significances unfortunately cannot be given throughout this paper (see Appendix 2 for more information).

3.4.1. General aspects questionnaire

Both groups of students (Law and Tourism) strongly agree that *pronouncing correctly in English is important*. Moreover, both groups believe it is not possible for one to speak well in English if one does not know how to pronounce isolated words well. Similarly, both groups agree that, generally speaking, Spanish EFL learners tend to have a Spanish accent when speaking in English. On the preference for native or non-native teachers, results indicate that both the Law and the Tourism subjects prefer the former, although the findings also suggest that Law students value native teachers slightly more than their counterparts in Tourism.

Some differences can also be observed if we compare average answers in each group. Law students, for example, agree that English pronunciation is difficult to some extent, whereas it does not seem to entail much difficulty for the students of Tourism. Another clear difference is that Tourism students do not believe it is important to speak English with a native accent, whilst those from Law agree (slightly) with this statement. Likewise, when asked whether they think a non-native English speaker is able to speak and pronounce like a native speaker, the views of the Tourism group are not as positive as those from the Law students.

Although both groups of students agree with the statement *it is possible for one to pronounce isolated words correctly without knowing how to speak English fluently*, the average score for the Tourism students was 6.5; this figure rises to 8.125 for the Law students.

A notable difference between the two groups is that Tourism students do not seem to mind listening to audio files featuring non-native speakers. Law students, on the other hand, appear to prefer listening activities with native speakers. Regarding intelligibility, the Law students clearly opted for the Standard British variety as the most understandable, whilst both the British and American standards were chosen quite frequently by the Tourism students. As can be seen in Appendix 1, the students were given the chance to choose other native varieties such as Irish, Australian, Canadian or South African; however, all opted for the two main standard models, probably because these varieties are the ones they are exposed to more frequently both in and outside the classroom (Walker, 2014).

	Law students	Tourism students
English pronunciation is difficult	6.625	4.6
Pronouncing English correctly is important	8.75	8.5
It is important to speak English with a native accent	5.43	2.83
It is possible to speak good English but not know how to pronounce the different words	3	2.5
It is possible for one to pronounce isolated words correctly without knowing how to speak English fluently	8.125	6.5
A non-native speaker of English is able to speak and pronounce like a native speaker	8.375	.16
Spaniards have a Spanish accent when speaking in English	7	8
I prefer having native English-speaking teachers than non-native ones	7.75	6.5
I prefer listening activities with native speakers than non-native ones	6.6	3.83
The native variety of English I understand best is...	BRITISH ENGLISH	BRITISH ENGLISH / AMERICAN ENGLISH

Table 13: Answers by the Law and Tourism students (mean values)

3.4.2. Students' opinions on different non-native accents

3.4.2.1. Polish speakers

Table 14 below shows the average scores by the Law and Tourism groups. Both Polish speakers were rated as having a good level of English pronunciation; moreover, both groups of students strongly believe the Polish person they heard can be classified as speaking well and fluently in English. Likewise, both groups of students stated that they could understand this speaker quite well.

Some differences across the two groups of students can also be observed. First, the Tourism students were generally much more critical with their Polish speaker when they were asked the extent to which this speaker's accent was similar to that of a native speaker, being more in agreement with the statement *this speaker has a strong foreign accent when speaking in English*. The Law students, on the other hand, rated the Polish speaker they heard as having an accent quite similar to a native one. Finally, whilst most of the Law students claimed they would be happy if they had an English teacher with this foreign accent, participants from the Tourism group were, on this occasion, far more critical.

	Law students	Tourism students
This speaker's accent is similar to that of a native speaker of English	7.5	2.66
This person has a strong foreign accent when speaking in English	5.6	7.83
This person speaks English well	8.83	8.3
This person pronounces English well	8	7.66
This person speaks fluently in English	8.3	8.2
I understand this person, they are intelligible	8	8
I would like to have a teacher of English with this accent	6.16	3

Table 14: Law and Tourism students' opinion on the Polish-accented speaker of English

The last question on the speaker-assessment questionnaires was *Where do you think this person is from?* As can be seen in Table 15, three Tourism students guessed that the speaker was indeed Polish. Strikingly, one student of Tourism (student number 2) rated this man as a native English speaker; moreover, two Law students and one Tourism one believed their Polish-accented speaker was Spanish.

	L1	L2	L3	L4	L5	L6
POLISH SPEAKER: GROUP A	SPANISH	----	EUROPEAN	SPANISH	---	RUSSIAN
	T1	T2	T3	T4	T5	T6
POLISH SPEAKER: GROUP B	EAST EUROPEAN	ENGLISH	POLISH	SPANISH	POLISH	POLISH / RUSSIAN

Table 15: Law and Tourism students' answers to *Where do you think this speaker is from?*

3.4.2.2. German speakers

Both the German speakers selected were rated as speaking and pronouncing well in English. Moreover, most students from both groups believed that these speakers were fluent. Both groups of students agreed with the statement *I would like to have a teacher of English with this accent*, although the average answers given on the Likert scale for this question do not exceed 6 out of 10 in either group.

As has been the case thus far, the findings also show some differences of opinion between the two participant groups. First, whilst the Law students believe their German speaker has an accent similar to that of a native speaker, when asked whether their German speaker could pass as a native speaker of English, the Tourism students clearly disagreed. Likewise, the Tourism students strongly believed that their German speaker

had a strong foreign accent, whereas the corresponding German speaker for the Law students was rated as not sounding too foreign-accented. Finally, both groups of students agreed that their German speaker was intelligible; nevertheless, the German accent was considered more intelligible by the Tourism students (an average score of 8.5 out of 10) than for the Law students (with an average of 6.25 out of 10 in their answers).

On this occasion, two Law students and only a single undergraduate in Tourism guessed the origin of the speaker (see Table 17).

	Law students	Tourism students
This speaker's accent is similar to a native English speaker's accent	7	4
This person has a strong foreign accent when speaking in English	4.25	7.16
This person speaks English well	8	8.16
This person pronounces English well	7.25	8
This person speaks English fluently	7.5	8
I understand this person, they are intelligible	6.25	8.5
I would like to have a teacher of English with this accent	6	5.6

Table 16: Law and Tourism students' opinion on the German-accented speaker of English

	L1	L2	L3	L4		
GERMAN SPEAKER:GROUP A	BRITISH	GERMAN	--	GERMAN		
	T1	T2	T3	T4	T5	T6
GERMAN SPEAKER:GROUP B	MEDITERRANEAN	SOUTH AMERICAN	--	SWEDISH	GERMAN	SPANISH

Table 17: Law and Tourism students' responses to *Where do you think this speaker is from?*

3.4.2.3. French speakers

The two French speakers in the assessments were both positively rated on the Likert scale in the questions *This person speaks English well* and *This person pronounces well in English*. The Law students strongly agreed that their French speaker spoke in the foreign language fluently; the corresponding French speaker assessed by the Tourism students was also rated as a fluent speaker, but the average response was not as high.

Once again, the Tourism students judged their (French) speaker as not sounding as a native speaker. For the Law students, on the other hand, the male French speaker they heard did remind them to a certain extent of a native English speaker. Despite these scores, the two groups agreed that their corresponding French speaker had a strong

foreign accent, although they also classified them as being intelligible and rather easy to understand. Finally, both groups were notably negative in response to the question of having an English teacher with a French accent.

	Law students	Tourism students
This speaker's accent is similar to a native English speaker's accent	7.83	3
This person has a strong foreign accent when speaking in English	6.33	7.6
This person speaks English well	7.33	7.5
This person pronounces English well	7.5	8.6
This person speaks English fluently	9	7.6
I understand this person, they are intelligible	6.3	7
I would like to have a teacher of English with this accent	4.6	2.3

Table 18: Law and Tourism students' opinion on the French-accented speaker of English

The French accent was the easiest for students to guess thus far. As can be seen in Table 19 below, a total of 8 students correctly perceived the origins of their French speakers (three Law and five Tourism students).

	L1	L2	L3	L4	L5	L6
FRENCH SPEAKER: GROUP A	FRENCH	--	FRENCH	FRENCH	--	BRITISH
	T1	T2	T3	T4	T5	T6
FRENCH SPEAKER: GROUP B	FRENCH	FRENCH	EAST EUROPEAN	FRENCH	FRENCH	FRENCH

Table 19: Law and Tourism students' responses to *Where do you think this speaker is from?*

3.4.2.4 Chinese speakers

When assessing the Chinese speakers, the Tourism students were (for the first time thus far) in relative agreement that they had an accent somewhat similar to a native English speaker. By contrast, the Law students on this occasion believe their Chinese speaker's accent could not be compared to that of a native speaker of English. For this latter group, the Chinese speaker had a strong foreign accent, whereas the Tourism students maintained a rather neutral position on this matter, neither clearly agreeing nor disagreeing. The Chinese speaker assessed by the students studying Tourism was rated as speaking English extremely well; moreover, respondents here believed that this woman spoke English very fluently and also had reasonably good pronunciation. The answers

given by the Law students to the previous three questions were also positive in terms of rating the speaker as a fluent speaker who speaks and pronounces English well; however, the average scores here were lower than those of the Tourism group (see Table 20).

Finally, both Chinese speakers were rated as being easy to understand, yet neither group of students was strongly in favour of having a teacher of English with a Chinese accent.

	Law students	Tourism students
This speaker's accent is similar to a native English speaker's accent	4.16	6.16
This person has a strong foreign accent when speaking English	7.3	5
This person speaks English well	6.6	9
This person pronounces English well	5.83	8.5
This person speaks English fluently	6	9
I understand this person, they are intelligible	7.3	7.5
I would like to have a teacher of English with this accent	4.5	5.33

Table 20: Law and Tourism students' opinion on the Chinese-accented speaker of English

Only two Tourism students guessed the origin of the Chinese speaker, although two more suggested a different Asian country. However, two Law students were extremely close since they thought this speaker was Japanese and two Tourism students mentioned their foreign speaker was from some Asian country.

	L1	L2	L3	L4	L5	L6
CHINESE SPEAKER: GROUP A	SPANISH	--	JAPANESE	SPANISH	--	JAPANESE
	T1	T2	T3	T4	T5	T6
CHINESE SPEAKER: GROUP B	SOUTH-EASTERN ASIA	ASIAN	BRITISH	CHINESE	DUTCH	CHINESE / JAPANESE

Table 21: Law and Tourism students' answers to the question *Where do you think this speaker is from?*

3.4.2.5. Spanish speakers

Both the Spanish speakers were considered as not having an accent similar to that of a native English speaker. Thus, the two groups agreed when asked whether *this person has a strong foreign accent when speaking English*. In general terms, however, both the Law

and Tourism students rated their Spanish speaker as speaking and pronouncing in English well and as speaking fluently. Nevertheless, the Tourism students had more positive views when answering the question *This person speaks English well*.

Finally, both groups affirmed that they understood their corresponding Spanish-accented speaker quite well but, surprisingly, although these speakers were intelligible for them they stated they would not like to have an English teacher with this non-native accent. On this occasion, the vast majority of students in both groups correctly identified the origin of these non-native speakers.

	Law students	Tourism students
This speaker's accent is similar to a native English speaker's accent	4.16	3
This person has a strong foreign accent when speaking English	7.6	6.83
This person speaks English well	6.83	8.33
This person pronounces English well	6.16	6.83
This person speaks English fluently	7.5	8.16
I understand this person, they are intelligible	8.16	8.66
I would like to have a teacher of English with this accent	4.6	2.83

Table 22: Law and Tourism students' opinion on the Spanish-accented speaker of English.

	L1	L2	L3	L4	L5	L6
SPANISH SPEAKER: GROUP A	MALLORCA	SPANISH	SPANISH	SPANISH	--	SPANISH
	T1	T2	T3	T4	T5	T6
SPANISH SPEAKER: GROUP B	SPANISH	SPANISH	SOUTHERN EUROPE	SPANISH	SPANISH	SPANISH

Table 23: Law and Tourism students' answers to *Where do you think this speaker is from?*

3.4.2.6. Students' opinions on each of the non-native accents. A comparison.

As shown in Tables 24 and Figure 1 below, the Spanish and French speakers received the lowest scores in three questions each. Specifically, the Spanish speakers were rated as: a) having the least similar accent to that of a native speaker, with an average score of 3.58, this in comparison, for instance, to the scores of 5.5 and 5.415 for the German and French speakers, respectively; b) having the strongest foreign accent when speaking in English; and, c) being the speakers judged lowest in terms of having good pronunciation in English. The French speakers, in turn, were considered to be: a) the ones who speak

English least well; b) the ones who are the least intelligible and thus the most difficult to understand; and, c) as having an accent that respondents would least like their English teacher to have. In addition to these findings, we might note that the Spanish non-native accents were also rated quite negatively in response to the statements *This person speaks English well* and *I would like to have a teacher of English with this accent*; the French speakers were also said to speak with a strong foreign accent in English. Curiously, and despite such negative assessments, the Spanish speakers were seen as having the non-native accent that participants best understood, whilst the French speakers were considered to have the best pronunciation in English and the ones who spoke most fluently.

At the other end of the scale, the German and Polish accented speakers were rated most positively on several questions. German speakers were considered to be: a) the speakers with the least strong foreign accent when speaking in English; b) the speakers that have the accent most similar to that of a native English speaker; and, c) the preferred non-native accent for an English teacher. The Polish speakers, in turn, were judged to be those who spoke English the best; furthermore, they scored quite highly on the items *This person pronounces English well*, *This person speaks English fluently*, and *I understand this person, they are intelligible*.

The Chinese speakers were not seen as the best or worst speakers in any of the questions, with the exception of the issue of fluency, in which they were rated as the least fluent speakers in English.

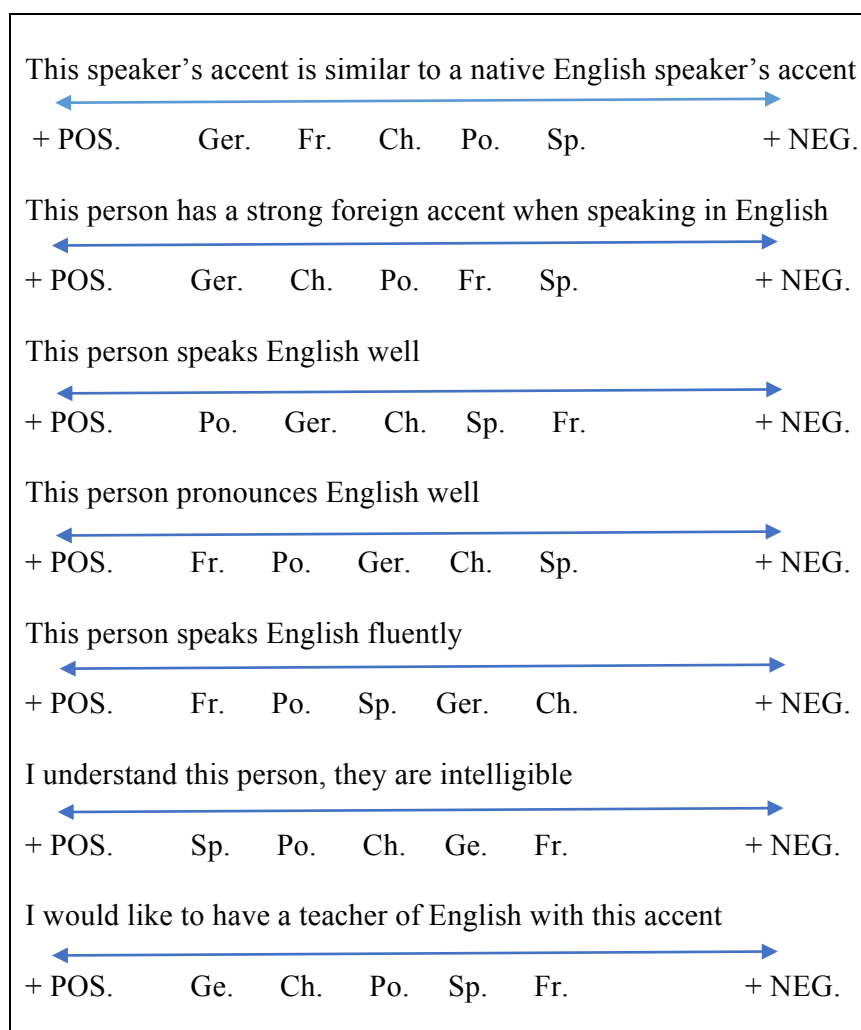


Figure 1: Positive-negative rating of non-native speakers for each question

MEAN SCORES FOR EACH QUESTION (GROUPS A AND B)	POLISH	GERMAN	FRENCH	CHINESE	SPANISH
This speaker's accent is similar to a native English speaker's accent	5.08	5.5	5.415	5.16	3.58
This person has a strong foreign accent when speaking in English	6.715	5.705	6.965	6.15	7.215
This person speaks English well	8.565	8.08	7.415	7.8	7.58
This person pronounces English well	7.83	7.625	8.05	7.165	6.495
This person speaks English fluently	8.25	7.75	8.3	7.5	7.83
I understand this person, they are intelligible	8	7.375	6.65	7.4	8.41
I would like to have a teacher of English with this accent	4.58	5.8	3.45	4.9	3.715

Table 24: Average scores by the Law and Tourism students in assessing each of the non-native speakers

4. Conclusions, teaching implications and topics for future research

Bayyurt and Sifakis (2017: 5) pointed out that “curricula and textbooks continue to serve a more traditional perspective of teaching and learning English” since Standard varieties of English are still prioritised. The general results obtained in the first study conducted here seem to perfectly illustrate this situation since ESP textbooks addressed to Law and Tourism students continue to have much more material containing native speakers as models. Non-native speakers, on the other hand, appear less frequently, and when they do, they tend to interact with native speakers; that is, there are very few tracks where only non-native speakers are heard.

Although this finding might be seen as somewhat negative, the general findings here are more positive than those of Kopperoinen (2011) (see section 1). Nevertheless, I consider that the level of exposure that ESP students have to non-native accents in the textbooks analysed is still not sufficient, especially in the case of the undergraduates of Tourism, since it is very likely that in their future professional life they will have to interact in English with many people from different language backgrounds and with different native languages. Such people may include: a) those from countries where English is an (additional) official language; b) second language speakers from countries where English plays an official role (Indians or South Africans, to mention just two); c) those from native English-speaking countries like Britain, Australia and the United States; and, d) non-native English speakers who typically speak in English to communicate when abroad, such as Germans, Polish, or Chinese; hence, all those speakers who use English as a Lingua Franca.

Despite this negative finding, it is worth mentioning that some of the textbooks analysed do contain quite a few examples of non-native speakers from different countries (mainly China, Japan, France, Germany, Italy, Poland and Spain), and hence they could be considered as good teaching materials, in that they give students opportunities to listen to people with these non-native accents. The most advantageous textbooks in this sense were *English for Legal Professionals* for Law students, and *English for Job Hunting* and *Tourism 3* for Tourism students.

As discussed in Section 2.1, the fact that the textbooks analysed were published a number of years ago may explain why some of them lack the presence of non-native speakers, since the phenomenon of ELF is relatively new. An analysis of the most recent textbooks on the market would therefore be interesting as a means of verifying the extent to which textbook content has shifted to embrace the new role of ELF, that is, whether more and more varied non-native accents are included, or whether native speaker models continue to be used much more frequently.

Several conclusions, as well as teaching implications, follow from the second study reported here. To begin with, the attitudes of both the Law and Tourism students towards learning spoken English are positive, in that they regard pronouncing correctly in English as important and believe that it is not possible to speak well in English if one does not know how to pronounce specific words well. Secondly, Law students seem to value native-accented over non-native ones, as seen in their preference for native teachers to non-native ones, the fact that they favour listening to native rather than non-native models of English in class, their belief that non-native English speakers can finally sound native-like over time, and that to a certain extent one should aim at speaking with a native accent when learning a foreign language. The Tourism students, on the other hand, tend to accept both native and non-native accents, and believe it is difficult for a non-native speaker to speak like a native speaker; for this reason, it is understandable that they rated speaking with a native accent in the foreign language as not necessary.

Some of the differences of opinion found between the two groups of students in the *General Aspects* questionnaire may be explained by the fact that, as explained above, the degree of exposure that students of Law and Tourism have to English is quite different.

Whilst the Law students in the study only have one optional ESP subject, Tourism students benefit from an ESP module in three out of the four years of their degree; moreover, it is common for Tourism students to work during the summer holidays, or even in the Spring onwards, in hotels, travel agencies and airports, places in which they are likely to interact with people who speak English as a lingua franca. Hence, the fact that English pronunciation is seen as difficult for Law students, but does not seem to be a problem for students of Tourism, can perhaps be explained by this different degree of exposure to the foreign language. Similarly, the general acceptance of non-native speakers identified in some of the questions answered by Tourism students might also be explained by this higher exposure to English and again by the fact that these students may already be used to talking in English to other non-native speakers due to their own professional experiences.

In general terms, although the opinions of the students who participated in this study vary from one non-native accent to another, they did not reject any of the non-native accents they were asked to assess. As might be expected in such a study on preferences, there are some accents which participating students preferred or understood better than others. Here, those speakers rated more positively in most of the Likert-scale questions were the German and the Polish ones; by contrast, the French and Spanish speakers received the lowest averages on some of the scores. Hence, as can be found in Jenkins (2007), the students believe some non-native accents sound better or are superior to others. In this case, the participants believe the native speakers of the two Germanic languages chosen (German and Polish) have a closer accent to that of a native speaker of English when communicating orally in the latter language; the native speakers of the two Romance languages chosen, Spanish and French, were rated as sounding less English native-like; hence, in general terms the Spanish and French speakers were rated as having an inferior English accent.

Regarding the Spanish speakers, and given that the ESP students who took part in the study were themselves Spanish, it may seem surprising that students assessed these as the ones who pronounced English least well, had the strongest foreign accent, and had accents which were the least similar to that of a native speaker. Moreover, most of the students said they would not want an English teacher with this foreign accent. Despite such negative results, the Spaniards were rated as the speakers who were the most intelligible. The latter finding is again understandable, in that Spanish students are most likely to understand other Spanish speakers, since they share the same features of accent, segmental pronunciation, etc; however, the fact that they affirmed they would not like an English teacher with this accent is surprising, as most of the EFL and ESP teachers in Spain are native Spanish speakers.

An area for future research would be to analyse the most recent ESP textbooks and compare results to those reported in the current study. In addition, it would be interesting to analyse the views of a larger number of students from different ESP disciplines regarding their perceptions on speakers with different non-native accents, since the present study was based on a small number of participants. A further interesting line of enquiry would be to conduct a study in which ESP students are asked to assess both native

and non-native speakers, in order to get a more involved picture of which accents are generally better understood.

All in all, then, I believe that this study has filled a gap in our understanding of the teaching of ESP in Spain. Several pedagogical implications can be drawn. First, some of the teaching materials currently used to teach ESP at the University of the Balearic Islands may be considered as appropriate, since a number of both native and non-native speakers are included in the audio materials. The textbooks which feature none or hardly any non-native speakers might also continue to be used, since they may contain useful material for teaching other language skills and areas, such as grammar or vocabulary; however, these materials should be supplemented with alternative audio materials in which non-native speakers are included, so that ESP students can get used to listening to speakers with different non-native accents speaking in English. Secondly, the fact that the Law students value native speakers over non-native ones so highly may simply be due to the few opportunities that they have on their university degree course for listening to non-native speakers of English; as I have noted, these students only study English during the last semester of their final year in the degree in Law, and thus may listen to far more native English speakers than non-native ones in class. A short-term solution would hence be for ESP teachers teaching Law students to try and select a greater amount of audio material in which native and non-native speakers interact. A more convincing solution would be to change some of the main teaching materials currently being used to teach Law students, or at least to combine them with others in which there is a higher presence of both native and non-native accents.

Finally, I believe that studies in which the opinions of students are taken into account, regarding aspects such as the methodology or teaching materials used in class, are crucial to developing a greater understanding of this area of language education. It would therefore be very beneficial, for both ESP students and instructors, if teachers were encouraged to ask their students about their preferences in this area. The issues of what activities students like or do not like, and why, what aspects of English they find most difficult, the reasons for this, are all likely to contribute to further improvements in educational outcomes.

Notes

1. The research described in this paper was funded by the Autonomous Government of Galicia (Grants numbers CN2012/81, GPC2015/004) and by the Spanish Ministry of Innovation and Industry (Grant number FFI2015-64057-P). These grants are hereby gratefully acknowledged.

2. Although only three textbooks were analysed in this group (*English for Legal Professionals*, *Absolute Legal English* and *International Legal English*), the latter contains 2 CDS, and each of these was analysed separately.

3. As mentioned above, the ESP subject taught to Tourism students in the third year, English III, revolves around job interviews and business meetings (with a chairperson, a secretary and department representatives); for this reason, the undergraduate students in Tourism were considered as perfect candidates to assess some tracks taken from the Tourism book analysed in

the previous study, with others extracted from the two Business English books, also analysed in that study.

Primary sources (Textbooks used for the analysis)

Callanan, Helen and Lynda Edwards (2010): *Absolute Legal English*. Peaslake: Delta Publishing.

Downes, Colm (2008): *English for Job Hunting*. Oxford: Oxford University Press.

Frost, Andrew (2009): *English for Professionals*. Oxford: Oxford University Press.

Krois Lindner, Amy (2008): *International Legal English*. Cambridge: Cambridge University Press.

Thomson, Kenneth (2007). *English for Meetings*. Oxford: Oxford University Press.

Walker, Robin and Keith Harding (2009). *Tourism 3*. Oxford: Oxford University Press.

References cited

Alseweed, Mohammad (2012): "University students' perceptions of the influence of native and non-native teachers". *English Language Teaching*, 5(12): 42-53.

Baker, Amanda (2011): *Pronunciation Pedagogy: Second Language Teacher Cognition and Practice*. Doctoral dissertation, Georgia State University.

Baayuurt, Yasemin and Nicos Sifakis (2017): "Foundations of an EIL-aware Teacher Education". In A. Matsuda, ed. *Preparing Teachers to Teach English as an International Language*. Bristol: Multilingual Matters, 3-18.

Buckingham, Louisa (2015): "Recognising English accents in the community: Omani students' accent preferences and perceptions of nativeness". *Journal of Multilingual and Multicultural Development*, 36(2): 182-197.

Celce, Marianne, Janet M. Goodwin and Donna Brinton (2010): *Teaching Pronunciation. A Course Book and Reference Guide* (2nd edition). Cambridge: Cambridge University Press.

Cheung, Chit (2010): "Native or non-native? Exploring Honk Kong students' perspectives". In G. Bota, H. Hargreaves, C.C. Lai and R. Rong, eds., *Papers from the Lancaster University Postgraduate Conference in Linguistics & Language Teaching*, 4. Lancaster: Department of Linguistics and English Language Lancaster University, 243-262.

Chun, Sun Young (2014): "EFL learners' beliefs about native and non-native English speaking teachers: perceived strengths, weakness, and preferences". *Journal of Multilingual and Multicultural Development*, 35(6): 563-579.

Dogancay-Aktuna, Seran and Joeal Hardman (2017): "A framework for incorporating English as an international language perspective into TESOL teacher education". In A. Matsuda, ed., *Preparing Teachers to Teach English as an International Language*. Bristol: Multilingual Matters, 19-32.

Dweik, Bader and Ekram Al-Barghouthi (2014): "Attitudes of Jordanian graduate students towards native and non-native English language teachers". *Journal of Education and Human Development*, 3(2): 627-646.

Grant, Linda (2014): "Prologue to the myths. What teachers need to know". In L. Grant, ed., *Pronunciation Myths. Applying Second Language Research to Classroom Teaching*. Michigan: University of Michigan Press, 1-33.

- Jaber, Maysa and Riyad Hussein (2011): "Native speakers' perception of non-native English speech". *English Language Teaching*, 4(4): 77-87.
- Jenkins, Jennifer (2007): *English as a Lingua Franca: Attitude and Identity*. Oxford: Oxford University Press.
- Karakas, Ali, Huseyin Uysal, Selcuk Bilgin and Talat Bulut (2016): "Turkish EFL learners' perceptions on native English-speaking teachers and non-native English-speaking teachers in higher education". *Novitas-ROYAL*, 10(2): 180-206.
- Kopperoinen, Anne (2011): "Accents of English as a Lingua Franca: A study of Finnish textbooks". *International Journal of Applied Linguistics*, 21(1): 71-93.
- Lasagabaster, David and Juan Manuel Sierra (2002): "University students' perceptions of native and non-native speaker teachers of English". *Language Awareness*, 11(2): 132-142.
- López-Jiménez, María Dolores (2009): "The treatment of vocabulary in EFL textbooks". *Estudios de Lingüística Aplicada (ELIA)*, 9: 59-81.
- Madrid, Daniel and M^a Luisa Pérez Cañado (2004): "Teacher and student preferences of native and nonnative foreign language teachers". *Porta Linguarum*, 2: 125-138.
- Marks Jonathan (2006): "Pronunciation in coursebooks again". *Speak Out!* 36: 33-35.
- Matsuda, Aya (2017): "Introduction". In A. Matsuda, ed., *Preparing Teachers to Teach English as an International Language*. Bristol: Multilingual Matters, xiii-xx.
- Matsura, Hiroko and Reiko Chiba (2014): "Students' attitudes toward unfamiliar English accents and their L1-speaking backgrounds". *商学論集 Journal*, 82(4): 3-13.
- Mauranen, Anna (2012): *Exploring ELF. Academic English Shaped by Non-Native Speakers*. Cambridge: Cambridge University Press.
- Mauranen, Anna, Ray Carey and Elina Ranta (2015): "New answers to familiar questions: English as a lingua franca". In D. Biber and R. Reppen, eds., *The Cambridge Handbook of English Corpus Linguistics*. Cambridge: Cambridge University Press, 381-400.
- Mauranen, Anna, Nina Hynninen and Elina Ranta (2016): "English as the academic lingua franca". In K. Hyland and P. Shaw, eds., *The Routledge Handbook of English for Academic Purpose*, London: Routledge, 44-56.
- Moinzadeh, Ahmad, Omid Rezaei and Salman Dezhara (2012): "The effect of non-native accent on Iranian EFL learners' listening comprehension, focusing on Persian accent of English". *Journal of Language Teaching and Research*, 3(5): 967-972.
- Motschenbacher, Heiko (2013): *New Perspectives on English as a European Lingua Franca*. Philadelphia: John Benjamins.
- Thamer, Zainab, Ain Nadzimah and Chan Swee Heng (2014): "Malaysian students' attitudes towards six varieties of accented speech in English". *Advances in Language and Literary Studies*, 5(5): 181-191.
- Seidlhofer, Barbara (2001): "Closing a conceptual gap: The case for a description of English as a lingua franca". *International Journal of Applied Linguistics*, 11(2): 133-158.
- Seidlhofer, Barbara (2011): *Understanding English as a Lingua Franca*. Oxford: Oxford University Press.
- Walker, Robin (2010): *Teaching the Pronunciation of English as a Lingua Franca*. Oxford: Oxford University Press.
- Walker, Robin (2014): "Horses for courses". *English Teaching Professional*, 91: 8-10.
- Walkinshaw, Ian and Duongthi Hoang (2014): "Native and non-native English language teachers: Students perceptions in Vietnam and Japan". *SAGE Open*, 4(2): 1-9.

Appendix 1: Questionnaires used in the study on ESP students' opinions of non-native accents.

A) General aspects questionnaire:

Choose the number which best suits your opinion in each of the following questions. 1 represents 'I totally disagree' and 10 'I totally agree'.

La pronunciación del inglés es difícil

1 2 3 4 5 6 7 8 9 10

Pronunciar bien en inglés es importante para que te entiendan

1 2 3 4 5 6 7 8 9 10

Es importante hablar inglés con un acento nativo

1 2 3 4 5 6 7 8 9 10

Prefiero tener profesores nativos de inglés que no nativos

1 2 3 4 5 6 7 8 9 10

Prefiero los *listenings* de hablantes nativos a los no-nativos

1 2 3 4 5 6 7 8 9 10

Se puede hablar bien en inglés sin saber pronunciar las palabras

1 2 3 4 5 6 7 8 9 10

Se pueden pronunciar bien palabras aisladas en inglés sin saber hablar con fluidez

1 2 3 4 5 6 7 8 9 10

Un hablante no-nativo de inglés puede llegar a hablar (y pronunciar) como uno nativo

1 2 3 4 5 6 7 8 9 10

Los españoles tienen acento español cuando hablan en inglés

1 2 3 4 5 6 7 8 9 10

La variedad del inglés que mejor entiendo es

- 1- Británico: de Inglaterra
- 2- Británico: de Irlanda
- 3- Británico: de Escocia
- 4- Inglés americano
- 5- Australiano/neozelandés
- 6- Sudafricano
- 7- Jamaicano
- 8- Otro

B) Questionnaires to assess the non-native speakers:

Choose the number which best suits your opinion in each of the following questions. 1 represents 'I totally disagree' and 10 'I totally agree'.

El acento de esta persona se asemeja al de un hablante nativo de inglés

1 2 3 4 5 6 7 8 9 10

Esta persona tiene un fuerte acento extranjero al hablar en inglés

1 2 3 4 5 6 7 8 9 10

Esta persona habla bien en inglés

1 2 3 4 5 6 7 8 9 10

Esta persona pronuncia bien en inglés

1 2 3 4 5 6 7 8 9 10

Esta persona habla con fluidez

1 2 3 4 5 6 7 8 9 10

Entiendo a esta persona, es inteligible

1 2 3 4 5 6 7 8 9 10

Me gustaría tener un profesor/a de inglés con este acento

1 2 3 4 5 6 7 8 9 10

¿De dónde crees que es esta persona? _____

Appendix 2: Statistical analyses for Tables 13, 14, 18, 20 and 22.

Item	Chi-square (X^2)	DF (degrees of freedom)	P value
Item 1	0.04	1	0.8415
Item 2	0.26	1	0.6101
Item 3	0.01	1	0.9203
Item 4	0.06	1	0.8065
Item 5	0	1	1
Item 6	0.03	1	0.8625
Item 7	0.17	1	0.6801
Item 8	0	1	1
Item 9	0.11	1	0.7401

Chi-square test used for the data gathered in Table 13.

Item	Chi-square (X^2)	DF (degrees of freedom)	P value
Item 1	0.58	1	0.4463
Item 2	0	1	1
Item 3	0.08	1	0.773
Item 4	0.11	1	0.7401
Item 5	0	1	0.8625
Item 6	0.03	1	0.6985
Item 7	0.13	1	0.184

Chi-square test used for the data gathered in Table 14.

Item	Chi-square (X^2)	DF (degrees of freedom)	P value
Item 1	0.51	1	0.4751
Item 2	0.02	1	0.8875
Item 3	0.15	1	0.6985
Item 4	0.04	1	0.8415
Item 5	0.14	1	0.7083
Item 6	0.07	1	0.7913
Item 7	0.07	1	0.7913

Chi-square test used for the data gathered in Table 18

Item	Chi-square (X^2)	DF (degrees of freedom)	P value
Item 1	0	1	1
Item 2	0	1	1
Item 3	0.06	1	0.8065
Item 4	0.01	1	0.9203
Item 5	0.02	1	0.8875
Item 6	0.02	1	0.8875
Item 7	0.05	1	0.8231

Chi-square test used for the data gathered in Table 20.

Item	Chi-square (X^2)	DF (degrees of freedom)	P value
Item 1	0.02	1	0.8875
Item 2	0.06	1	0.8065
Item 3	0.01	1	0.9203
Item 4	0.07	1	0.7913
Item 5	0.08	1	0.7773
Item 6	0.09	1	0.7642
Item 7	0.01	1	0.9203

Chi-square test used for the data gathered in Table 22.

Rhoticity in Chinese English: An experimental investigation on the realization of the variant (r) in an Expanding Circle variety

Zeyu Li
University of Münster
zeyu.li@uni-muenster.de

Barış Kabak
University of Würzburg
baris.kabak@uni-wuerzburg.de

ABSTRACT

The realization of postvocalic /r/ has been frequently examined in both diachronic and synchronic research on world Englishes, showing a multitude of linguistic and extra-linguistic factors to modulate the degree of rhoticity. Since rhoticity is one of the most important indices of variation across Englishes, it forms an instructive phonological marker to investigate the dynamics of norm formation in emerging varieties. While the Inner and Outer Circle varieties have been extensively studied, there is fairly little research on the variable realization of postvocalic /r/ in the Expanding Circle Englishes. Here, we fill this gap with a study on the degree of rhoticity by highly proficient users of an EFL variety emerging in China, college English teachers, who are pertinent norm providers for EFL learners. We provide a multivariate analysis of phonological and sociolinguistic factors conditioning the degree of rhoticity in Chinese English on the basis of speech production data from 13 participants. Results show that Chinese English is best categorized as marginally rhotic. Concerning the patterning of phonological variables, it aligns more with Inner Circle than Outer Circle Englishes, albeit with significant inter- and intra-speaker variability. We discuss the competing roles of norm orientation, substrate influence, and other relevant variables therein.

Keywords: Expanding Circle, Phonology, Rhoticity, EFL varieties of English



1. Introduction

Due to their ambivalent phonetic properties, rhotic sounds are one of the most variable segments across the languages of the world, not only in their actual phonetic realization but also in their distribution within different word positions and phonological contexts. The abundance of empirical work on the realization of postvocalic /r/¹ in English is a clear manifestation of the high degree of variability that this phoneme exhibits within individual speakers as well as across different speaker groups. Indeed, the presence or absence of postvocalic /r/ has long been considered one of the most salient features that robustly index dialectal differences in World Englishes (Trudgill and Hannah, 1994: 5-6). Broadly, accents of English that produce /r/ in syllable rhymes (also variably referred to as postvocalic /r/, coda /r/, or non-prevocalic /r/) are considered to be rhotic, whereas those that do not realize /r/ in this context are non-rhotic (Giegerich, 1992: 63). Accordingly, while non-rhotic accents pronounce /r/ only in syllable onsets as in *red* or *break*, rhotic accents would also produce the same category in the postvocalic position as in *here* and *hurt*. Despite this coarse distinction, rhoticity is known to be a gradient phenomenon. The classic work of Labov (1966) on New York City English, for instance, has shown that speakers go beyond the binary rhotic-nonrhotic distinction and instead show gradient patterns in their realization of postvocalic /r/. Along with grammar-external factors such as age, ethnicity, and socioeconomic background, Labov also demonstrated that grammar-internal phenomena such as stress and the following phonological environment crucially modulate the degree of rhoticity.

Several replications of Labov's groundbreaking study and its extensions to other varieties have not only yielded very similar patterns but also found additional factors such as the quality of the preceding vowel to be a predictor of the rate of rhoticity (see, for example, Fowler, 1986; Mather, 2012; Becker, 2009, 2014 for New York City English; Nagy and Irwin, 2007, 2010 for Boston English, and Feagin, 1990 for Alabama English). Furthermore, Labov's work showed that non-rhoticity as a historically prestigious variant in New York City has lost its prestige over time, adding lifespan dynamics to this variant. Such a dynamic approach has led to further work on other varieties of English, such as New Zealand English, where the loss of postvocalic /r/ was shown to be a gradual weakening process, but crucially progressing independently from the emergence of /r/-sandhi phenomena such as /r/-linking and /r/-intrusion (Hay and Clendon, 2012).

Apart from its importance in the field of sociolinguistics, the variant (r)² constitutes one of the most salient criteria that distinguish types of English around the globe (Trudgill and Hannah, 1994: 5-6) and its distribution in postvocalic position is considered to be "the most important phonotactic difference among accents of English" (Melchers and Shaw, 2003: 19). Thus, it may be possible to establish a phonological typology of World Englishes based on rhoticity if the (r) variant in each English variety were thoroughly studied. To date, rhoticity within Inner Circle Englishes (Kachru, 1982, 1985) has been well investigated in comparison to nonnative varieties of English. The non-rhotic varieties are found in most of England, Australia and New Zealand, whereas varieties elsewhere such as those found in Scotland, Ireland and most of the USA and Canada are predominantly rhotic (e.g., Giegerich, 1992: 63; Sundkvist and Gao, 2016: 43). A number

of studies have documented a change toward more rhotic realizations in the Inner Circle areas, where the variant (r) has been historically unstable. For instance, a great deal of quantitative sociolinguistic research investigating (r) usage has shown that the historically prestigious non-rhoticity in various parts of the United States such as New York (Labov, 1966, 1972), Boston and New England (Irwin and Nagy, 2007), Philadelphia (Miller, 1998), and Upper Valley (Villard, 2009) has gradually lost its high social valuation.

While the variant (r) has attracted a great deal of interest also from the perspective of Outer Circle Englishes, as we will review below, there is very little on its nature and dynamics in Expanding Circle Englishes. We consider the latter types of English to constitute incipient non-native varieties that are on their way to establish their unique sociocultural stature and perpetuate novel norms within certain regions of the world. This paper aims to fill that gap with an empirical study on the degree of rhoticity in highly proficient nonnative users of an English variety emerging in China, which we will heretofore refer to as Chinese English.³ In particular, we will examine the nature and patterning of the variant (r) in the speech of English teachers, whom we assume to be one of the most pertinent norm providers for the learners of this variety. We will provide a multivariate analysis of phonological factors conditioning the degree of rhoticity in Chinese English and also explore the competing influence of extra-linguistic factors such as norm orientation, substrate influence, and other sociolinguistic variables therein. Our results take the debate on the dynamism of World Englishes to the dimension of *growth*. In its canonical sense, growth may refer to an increase in number, size, magnitude or value, and in the linguistic context we also consider it to be a result of a process of acquisition, reinterpretation and extension of existing linguistic schemes and patterns that may yield seemingly “novel” patterns of behavior. These novel patterns are however very likely to be conditioned by factors that underlie language acquisition processes and language change. To that end, we will discuss the consequences of our findings in the framework of the Dynamic Equational Model (Kabak, to appear), which views sound systems as empirical constructs at any *point* and *type* of language evolution (e.g., first/second language acquisition, language attrition, diachronic and synchronic variation, in the emergence of varieties, etc.) and aims to unite these different types of sound evolution as being governed by similar laws, thus considering them equally complex and natural.

The paper is organized in the following way. In Section 2, we will review findings from previous research concerning the nature of rhoticity in Inner Circle (2.1), Outer Circle (2.2.), and Expanding Circle Englishes (2.3) in an attempt to lay ground for the present study. Section 3 provides a phonological characterization of rhoticity in the substrate language in question, Mandarin Chinese. We present the design features of our study, data elicitation procedures, and our experimental conditions in Section 4, followed by our multivariate analysis of the results in Section 5. Section 6 unpacks the linguistic and extra-linguistic variables that turned out to be significant factors in the present study against the background of previous studies on the variant (r). Section 7 concludes with the general implications of the findings for the evolution of sound systems with some suggestions for future research.

2. The variant (r) in Englishes around the globe

We adopt the framework of Kachru (1982, 1985) in our classification of World Englishes, whose approach resonates with our interest in the developmental aspects and the genesis of new norms in localized varieties of English. While this framework may fail to reflect some of the dynamic changes that have taken place especially in Inner Circle Englishes through time (see *The Dynamic Model* of Schneider 2003, 2007), its social, political and practical implications on the conceptualization of growing EFL users around the globe remain well-founded. According to Kachru, the first circle is called “the inner circle”, which is located at the center of the concentric model and contains countries where English is spoken as native language, such as America, Britain, Australia, New Zealand, and Canada. The second group is “the outer circle”. It surrounds the inner circle and comprises places where English is used as a second language such as Singapore, India, Tanzania and Kenya, etc. The biggest group is “the expanding circle”, which includes countries such as Korea, China and Germany that use English as a foreign language. Given that most systematic studies on rhoticity have so far focused on Inner Circle Englishes, some of which we outlined above, below we focus on Outer and Expanding Circle Englishes. Needless to say, there is also a considerable amount of work on the variant (r) from the perspective of social and ethnic variation, focusing for example on African American English in different regions of the United States (see Thomas, 2007: 453-454 for a review). Furthermore, a few other factors have recently been shown to predict the degree of rhoticity in varieties of English such as the topic of conversation and exposure to a (non)rhotic variety (Love and Walker, 2012) as well as style shift (Eberhard and Downs, 2015). Since our primary aim here is to put the variant (r) on a more global perspective and to study the genesis of phonological repertoires in newer forms of English in regions with greater masses of EFL learners, we will not cover socioethnically and pragmatically modulated rhoticity here.

2.1. Rhoticity in Inner Circle Englishes

In order to pinpoint those factors that are time and again deemed important in the investigation of the variable nature of rhoticity, here we will discuss recent research on two Inner Circle Englishes, one from the United States and the other from the United Kingdom.

Nagy and Irwin (2010) investigated the realization of (r) in English spoken in Boston and two New Hampshire (NH) towns and described the effects of both internal and external variables on (r) usage by 55 participants. The study has shown that the preceding vowel and variation in geography, ethnicity, and age significantly influence the presence or absence of (r) in Boston and the two NH towns. It has also been suggested that English in these areas is gradually gaining rhoticity as younger speakers were found to be more rhotic than older speakers. We consider Nagy and Irwin (2010) representative and comprehensive in its kind because it additionally provides an in-depth comparison of studies of (r) in North American English and introduces a method to compare the quantitative amount of variation between communities. More precisely, Nagy and Irwin

explored certain factors that have repeatedly emerged as relevant to all rhoticity variation studies in the United States and turned to a comparison of variationist studies of (r) across North America to report on the relative universality and dissimilarity of these different factors.

Piercy (2012) presented a similar line of research concerning rhoticity in Dorset English, a variety from the Southwest of England that is originally rhotic. This study has focused on, among other linguistic factors, extra-linguistic factors such as age and gender that are considered to condition (r). An experiment based on 24 native speakers from Dorset showed that this historically rhotic variety in England has been gradually losing its rhoticity as evidenced by the decrease in the frequency of (r) tokens in both real and apparent time. While Piercy found a strong correlation between age and (r) in apparent time, a real time point of comparison to the Survey of English Dialects (SED) showed a steep decline of the overall (r) frequency from 97% in the SED to 29% in her study. Furthermore, the preceding vowel, word context, stress and lexical frequency constituted significant linguistic factors controlling (r) in her study. Piercy (2012: 83-84) further provided a transatlantic cross-dialectal comparison to see whether varieties losing or gaining rhoticity are influenced by the universal linguistic constraints. In particular, the results from the Dorset English study were compared with those from Nagy and Irwin (2010), showing that there are similarities in the linguistic contexts that determine the presence/ absence of /r/ irrespective of whether an Inner Circle variety is gaining or losing rhoticity (Piercy, 2012: 85).

The following review of studies will only be cursory to view the variant (r) from a more global and typological perspective. We will recapitulate some of the factors that were consistently shown to be relevant for the variant (r) in our methodology section in order to motivate the design features of our study and highlight some predictions.

2.2. Rhoticity in Outer Circle Englishes

With regard to the Outer Circle (Kachru, 1982, 1985), it is more challenging to give an explicit statement on the nature and extent of rhoticity (Melchers and Shaw, 2003: 131). In general, varieties that are traditionally considered to be predominantly rhotic are Indian English (Wells, 1982, but see below), Pakistani English (Mahboob and Ahmar, 2004), Maltese English (Bonnici, 2010), and Philippine English (Llamzon, 1997) while other regions are considered to be mostly non-rhotic, including Hong Kong (Setter et al., 2010), Malaysia and Singapore (Deterding, 2007) in Asia, East African countries like Kenya and Tanzania (Schmied, 2006), and West African countries such as Nigeria (Gut, 2004). However, Melchers and Shaw (2003: 131) point out that there has been little agreement on whether a particular variety in the Outer Circle is rhotic or not, since the depiction of a particular English variety on the grounds of rhoticity could be sensitive to a variety of endogenous and exogenous factors, as shown by previous research. For example, Salbrina and Deterding (2010) reported that Singapore English was mostly non-rhotic, with only 8.3% of their test tokens showing rhoticity. In contrast, the earlier work by Tan and Gupta (1992) revealed some degree of rhoticity in Singapore English, and further demonstrated that the /r/-fullness is a prestigious feature for some Singaporeans. Poedjosoedarmo

(2000), as cited in Tan (2012: 2), also suggested that American English-like characteristics were found in Singapore English; however, very few realizations of (r) have been noted in the database. In a similar vein, Tan (2012: 12-13) argued that speakers of higher education levels and socioeconomic status have a tendency to produce more (r)s (55.9%), compared to speakers from lower social class and education background (10.32%). Tan further suggested that her results seem to follow the classic Labovian hypothesis (Labov 1966) on the use of (r) by speakers from different social stratifications in New York City (Labov, 1966). According to Tan, this kind of ambivalence with regard to the degree of rhoticity in Southeast Asian Englishes is universal.

Considering the complex sociolinguistic circumstances surrounding the Outer Circle, Melchers and Shaw (2003: 131) have argued that there are five major factors exerting potential influence on the rhoticity in postcolonial Englishes. These are: substrate influence, education, source language, model, and endonormativity. To begin with, the substrate influence factor predicts that if sequences such as *VrC* or *Vr#* do not exist in the substrate language of a particular country, this will weigh against rhoticity. In contrast, substrate languages that allow such structures (e.g. Indian languages and Malay) would favour rhoticity. As for education, Melchers and Shaw (2003: 131) indicated that school learning would tend to produce pronunciation patterns reflecting the orthographic representation of sounds in spelling. In terms of endonormativity, English varieties introduced by speakers from England are expected to show a tendency to be non-rhotic while those instituted by Americans or speakers from Scotland and Ireland will tend to be rhotic. Likewise, the pronunciation patterns of individuals that act as speech models such as teachers and speakers in the local mass media would be very likely to influence the outcome. The local attitude, in addition, also plays an important role in determining the degree of rhoticity. It is claimed that if the local attitude in a particular region is highly endonormative, speakers would be more likely to speak in the local way conditioned by the above-mentioned factors, whereas a relatively exonormative attitude would lead to an accent that is closer to the dominant American pattern.

A number of studies on rhoticity in Outer Circle Englishes investigated the complex interplay of the above-mentioned five factors as well as other determinants. For instance, Chand's (2010) research on the usage of (r) in urban Indian English illustrated a diachronic change in the degree of rhoticity across three generations. Despite earlier claims that Indian English is a rhotic dialect (Wells, 1982: 629), only the oldest generation in Chand's data was found to be mostly rhotic. While speakers of the working generation preferred to speak non-rhotic English, the youngest generation tended to be more rhotic at home yet non-rhotic in formal situations. Several sociolinguistic factors have been shown to be influential in shaping the development of rhoticity in this variety such as distinctive postcolonial history, governmental policy, joint family system, change of instruction languages in school, linguistic attitude and various media inputs. Chand (2010: 22-25) has suggested that the oldest generation in the pre-partition India appeared to be more rhotic due to the influence from the rhotic substrate language and multiple prestige targets. She showed that the younger working generation, in contrast, tended to be non-rhotic, since the target pronunciation in India was narrowed to Received

Pronunciation (RP) after Partition. Non-rhotic accent was also used as the official instruction language in school. Interestingly, the current youth were found to be mostly rhotic and speak like their grandparents at home owing to the joint family system but switched to RP in the formal speech context like their parents.

Similar research has been conducted by Bonnici (2010) on Maltese English, who drew her data from sociolinguistic interviews with 32 (16 male and 16 female) English dominant or balanced Maltese-English bilinguals from Malta, ranging in age from 18 to 81. The results of this study have demonstrated an apparent time change in progress, with a move from extremely /r/-full behavior to /r/-lessness and back to increased /r/-fullness. This tendency in Maltese English is found to be subject to multiple sociolinguistic constraints such as dramatic shifts in speakers' exposure to rhotic varieties, local stigmatization of salient features of RP, as well as the rhotic structure of the substrate language, Maltese.

Hartmann and Zerbian (2010) explored rhoticity in the speech of black speakers of South African English, investigating whether the degree of rhoticity in South African English is sensitive to gender and socioeconomic stratification. Using various elicitation tasks (e.g., reading a disclaimer, answering simple questions with "better" or "worse", etc.), Hartmann and Zerbian conducted 39 interviews with both male and female native black South African participants. The two gender groups were further classified into more and less affluent categories (see Hartmann and Zerbian, 2010: 138-139 for the criteria). It was shown that rhoticity in Black South African English is predominantly a female phenomenon. While the degree of rhoticity was the highest among rich South African female participants (38%), "the less affluent females showed rhoticity in 18% of the cases on average" (p. 140). Interestingly, while 90% of all the affluent females in their sample displayed some degree of rhoticity, the mean rate of rhoticity was 8% among the affluent males, with only 40% of such participants showing a small amount of rhoticity (which included even the use of linking-r). However, none of the *less* affluent male participants showed any /r/ presence. This result was attributed to the fact that there is a covert prestige for African men to use their African languages. Accordingly, Hartmann and Zerbian (2010: 139-140) have concluded that their research provides evidence that both gender and affluence determine the degree of rhoticity in Black South African English and suggested that black female speakers are more progressive in their language use than black male speakers, as mirrored, for instance, by their more frequent use of English with their peers and family (p. 144). It should be noted however that their study overlooks the phonological aspects of rhoticity within this historically non-rhotic English variety.

2.3. Rhoticity in Expanding Circle Englishes

In contrast to other types of English, the variant (r) in Expanding Circle Englishes is severely understudied. At the point of writing this paper, we encountered only two relevant studies, one on Korean English (Kang, 2013) and another on Yunnan English (Sundkvist and Gao, 2016), both of which we will unpack below.

Based on an empirical study with a test sample of 30 college students from Korea, Kang (2013) reported that the Korean variety of English is more rhotic and closer to

American English. Meanwhile, both internal and external constraints controlling the variable (r) were explored. Kang's results showed that the external variables outweigh internal factors and exert a greater impact on rhoticity in Korean English (e.g., speech style, English-speaking nation living experience, proficiency level, and sex). Two limitations about this research were reported by Kang. First, the experiment in this study was exclusively based on reading speech, with the passage reading task as the task to elicit informal speech style and the word-list reading task to elicit less informal speech. Variation in speech styles, however, could not be well demonstrated by the two tasks that relied on reading. What is necessary here is spontaneous speech to uncover any potentially task-induced variation. The other limitation, according to Kang, concerns the age of speakers. Since only young college students were recruited as informants, age could not be examined as a potential external factor. Thus, Kang suggested that older speakers should be employed in the future research.

Sundkvist and Gao's (2016) study on 'Yunnan English' also explores the nature of rhoticity in Expanding Circle Englishes, in which 8 college students from Yunnan province in China were presented with various tasks representing three different speech styles. Their results showed that the most informal interview task displayed the lowest degree of (r) presence (38.5%), followed by the reading task with a slightly higher percentage (45.1%). The highest rate of rhoticity (64.9%), however, came from the most formal questionnaire task. Accordingly, Sundkvist and Gao (2016: 55) concluded that the variation in speech styles is influential to rhoticity in this nonnative variety due to various levels of formality and attention triggered by the tasks. Their work, according to Sundkvist and Gao (2016), has added to the limited empirical work on the rhoticity of English in the Expanding Circle. Furthermore, it has extended the inquiry into style-triggered variation, on which little work has so far been published. However, various potential linguistic and non-linguistic constraints on rhoticity in the Yunnan variety of Chinese English were not systematically controlled for. Additionally, the researchers do not discuss the potential contribution of L1 rhoticity. As we will see below, the Yunnan variety of Mandarin Chinese is largely non-rhotic just like many Mandarin varieties spoken in the South. Since all of the 8 participants spoke this dialect, it is not possible to tell whether Chinese English speakers with more rhotic backgrounds (such as those from North China) would also exhibit this type of variation that is commensurate with different degrees of formality. Finally, it should be noted that, as in Kang's (2013) study, this study drew its findings from test groups comprised of students.

2.4. Motivating the Present Study

While studies on EFL learner groups may echo properties of incipient EFL learner varieties, we find it necessary to investigate an Expanding Circle English from the point of view of the formation and perpetuation of endonormative practices in the Expanding Circle. To that end, it is essential to investigate a group of EFL users comprised of highly proficient speakers who use English regularly in their daily lives and who serve as models for incoming users of English in their locality—*English teachers*.⁴ Furthermore, since it remains to be a challenge to determine whether an English variety is rhotic or not due to

the multitude and complexity of factors, studies that take into account both language-external and language-internal variables are necessary to gain a comprehensive understanding of rhoticity in Expanding Circle Englishes. Given the fact that there are very few studies on this phonological variable within this relatively understudied type of English, here we aim to examine the degree of rhoticity in one of the most prominent varieties in the Expanding Circle, Chinese English, with special emphasis on both the endogenous and exogenous constraints affecting the variable realization of (r). In particular, we test whether Chinese English is rhotic or non-rhotic in light of the phonological constraints that were previously reported to affect variation in (r) in Inner Circle Englishes. We also examine the role of extra-linguistic factors such as target norm and L1 rhoticity influencing Chinese English rhoticity.

3. Substrate influence: Rhoticity in Mandarin Chinese

Rhoticity in the substrate language of a particular English variety has been considered a significant factor influencing the variable (r) in English (Melchers and Shaw, 2003), which makes it necessary for us to discuss the nature of rhoticity in the relevant substrate language of our Chinese English group, Mandarin Chinese.

Due to the severe restrictions on the type of consonantal segment that can occur in the postvocalic position (Ching 1973, Hansen 2001), postvocalic rhoticity in Mandarin Chinese is mainly attributed to the use of the suffix /ʔ/ (*/l* in simple Chinese character), which is a distinguishing characteristic of the Beijing dialect that is widely used in China. According to Duanmu (2007: 223-224), many Chinese dialects in the Mandarin family have this diminutive rhotic retroflex suffix, which is mostly used in North China. Mandarin spoken in the South of China, in contrast, is predominantly non-rhotic and most Southern Chinese dialects such as Cantonese and Hokkien also do not have rhotic accents at all.

The analysis of /ʔ/-colored syllables in Mandarin is not straightforward. Duanmu (2007: 224) assumes that the rhotic suffix is underlyingly specified only for the features Coronal and [+retroflex], which then merges with the host syllable that it attaches to. Two main merging processes are reported. First, if the sounds in the host rhyme are incompatible with [+retroflex], such as /n/, /i/, and /y/, they would either be replaced by /ʔ/ or pushed into the onset. For instance, when the rhotic suffix attaches to [p^han] ‘plate’, it is realized as [p^haʔ], whereby /n/ is replaced by /ʔ/. When the same suffix is attached to [y] ‘fish’, the rhyme is pushed into the onset, yielding [ʏʔ]. Elsewhere the suffix is simply added to the coda position or merged with the syllable it attaches to (e.g., [x^waa] ‘flower’ → [x^waʔ]). In all of these cases, the stems affixed by the rhotic suffix are rendered rhotic.

According to Chen (1999: 34-41), the rhotic suffix in Mandarin Chinese is mainly used in one of the following three situations. Firstly, it may serve to semantically

differentiate words (e.g., [paɪmiɛn] ‘flour’ vs. [paɪmiɛə̃] ‘heroin’). Secondly, Chen argues that many Mandarin words habitually occur with the rhotic suffix in North China, while their non-rhotic counterparts, despite acceptable, would sound unnatural to rhotic speakers. For example, [x^waa], which means ‘flower’, is usually produced as [x^waə̃] by rhotic Mandarin speakers while being realized primarily as [x^waa] by the non-rhotic speakers from the South China. Another important fact about the rhotic suffix in Mandarin involves the stylistic effect it achieves. In particular, Chen argues that the rhoticized form in Mandarin is typically confined to informal situations, while the corresponding unrhoticized form is used more frequently in formal speech contexts. This style-dependent variation can be taken to suggest that the degree of rhoticity in English for a particular rhotic Mandarin speaker is likely to vary significantly across different speech styles. For example, casual speech produced by a rhotic Mandarin speaker is likely to be colored by an abundance of rhotic sounds, increasing the likelihood of /r/-fullness, whereas this rate may decline in the same speaker when delivering a speech in a more formal situation. We take into account this potential stylistic effect in our experiment.

4. Methodology

4.1. Participants

Thirteen speakers (8 female, 5 male) from China were recruited for the present study. All of them were college English teachers at the North China Institute of Science and Technology (NCIST), a university that is located in the east of Beijing. We elicited participant-internal and demographic data by way of a questionnaire (see the appendix), which revealed the following characteristics. Despite its location in North China, informants of the present study were of various provinces of origin in China. All were native Mandarin speakers, aged between 31 and 59. Ten of them reported to speak particular Chinese dialects besides Mandarin due to their place of origin (the remaining 3 did not self-report any dialectal background). In addition, all of them used English as their first foreign language, with the knowledge of different additional foreign languages with varying levels of proficiency. It is worth noting that none of their additional foreign languages is rhotic. While 7 of them reported to use General American English as their target norm, 6 speakers preferred Received Pronunciation (RP), which provided us with a balanced set insofar as testing the contribution of L1 rhoticity. In addition, none of them had been staying abroad in an English-speaking country for more than a month. All the participants were tested at NCIST in China. The background statistics of the informant set is summarized in Table 1.

Speaker	Gender	Age	Place of origin	Native language	Chinese dialects	First foreign language	Other foreign language	Experiences abroad	Target Norm	English Media Exposure
1	F	45	Heilongjiang	Mandarin	none	English	French	England	BE	both
2	F	39	Liaoning	Mandarin	Dongbei	English	Japanese	US	AE	both
3	M	50	Hebei	Mandarin	Hebei	English	French	US	AE	both
4	F	31	Hebei	Mandarin	Jin	English	German	England	AE	AE
5	F	34	Hebei	Mandarin	Hebei	English	French	none	AE	AE
6	M	47	Hubei	Mandarin	Hubei	English	Ruassian	US, Austrilia	BE	both
7	F	35	Jiangxi	Mandarin	Jiangxi	English	Ruassian	none	BE	both
8	F	59	Jiangsu	Mandarin	Xuzhou	English	Japanese	US, Austrilia	BE	both
9	M	40	Hunan	Mandarin	Hunan	English	French	none	BE	BE
10	F	35	Shandong	Mandarin	Shandong	English	French	none	AE	both
11	M	44	Hebei	Mandarin	none	English	Japanese	Austrilia, New Zealand	AE	AE
12	F	47	Hebei	Mandarin	none	English	German	none	AE	both
13	M	38	Shandong	Mandarin	Henan	English	Japanese	Austrilia	BE	BE

F: female; M: male; BE: British English; AE: American English

Table 1: Background Statistics of Participants

4.2. Data collection

The elicitation of speech data in the present study consisted of three tasks: 1) a reading aloud task, where the participants were given a short story in English, 2) a free speech task in English, and 3) a reading aloud task in Mandarin Chinese, which were administrated in this particular order for all participants. All participants were recorded in an acoustic laboratory at the NCIST individually using a high-quality microphone

(NUEMANN U87). The speech was digitalized onto a computer (Apple Mac Pro MC561CH/A) at a sampling rate of 48000 Hz.

The reading task in English enabled us to collect fixed speech data by controlling for the speech style, test tokens, and linguistic environments. For this, we used a short English text, *The Wolf Passage*, an adaptation of Aesop's fable "The Boy who Cried Wolf" by Deterding (2006). This passage was specifically designed to elicit phonetic data in English and is claimed to contain clear tokens of all the vowels and consonants of English (Salbrina and Deterding, 2010). Besides, it includes certain low frequency phonemes that were found to be particularly relevant in the research of Asian Englishes (Sundkvist and Gao, 2016). We employed two additional paragraphs that are thematically related to the fable to generate a balanced design in order to test various phonological conditions that have been previously deemed important (the quality of the preceding vowels and the type of syllable structure). Altogether the reading task yielded 36 test tokens, testing 6 different vowel categories and 2 different syllable structures (Table 2).

	NURSE	LETTER	NEAR	SQUARE	NORTH	START
/r/ in Coda position	<i>emergency</i> <i>version</i> <i>perfect</i>	<i>pleasure</i> <i>however</i> <i>louder</i>	<i>near</i> <i>nearby</i> <i>appear</i>	<i>there</i> <i>air</i> <i>where</i>	<i>more</i> <i>unfortunately</i> <i>before</i>	<i>farmers</i> <i>hearted</i> <i>far</i>
/r/ in Consonant Clusters	<i>heard</i> <i>concern</i> <i>third</i>	<i>southern</i> <i>shepherd</i> <i>villagers</i>	<i>weird</i> <i>appeared</i> <i>years</i>	<i>scared</i> <i>impaired</i> <i>repaired</i>	<i>short</i> <i>course</i> <i>warm</i>	<i>dark</i> <i>alarm</i> <i>harmful</i>

Table 2: Stratification of test tokens by preceding vowel and syllable structure in reading task

After reading aloud the passage, the participants were given a free speech task, in which an approximately five-minute talk around the suggested themes was expected. In particular, they were asked to select one or two topics from the following six options: food, sports, Internet, weather, marriage, music, and festival, and then to talk about it or answer relevant questions for approximately five minutes. Given that a number of studies have shown that variation in speech styles significantly influence the performance of English speakers (e.g. Labov, 1972; Irwin and Nagy, 2010; Sundkvist and Gao, 2016) and that the rhotic suffix in Mandarin is assumed to be more frequent in colloquial language, we asked whether this type of a task would reveal differences in the amount of rhoticity in comparison to a more formal reading task.

In order to examine speakers' degree of rhoticity in their native language, which is imperative to the investigation of potential cross-language influence effects on the realization of (r), we additionally asked speakers to read aloud a Mandarin sentence that contained 10 common words that are typically rendered with rhotic vowels in rhotic varieties of Mandarin. The rhotic suffix 'ㄝ' in Mandarin Chinese was not presented in the text to avoid the potential impact of orthography. The assumption was that rhotic Mandarin speakers were expected to produce rhotic sounds at the end of these words habitually, regardless of whether the rhotic suffixes were in the text or not.

Following the recording session, informants filled out a demographic questionnaire, in which information about their personal background and the preferred English variety was collected (see the appendix). The final part of the questionnaire involved 10 statements about RP and American English (e.g., “British English /r/-less accent conveys a sense of arrogance.”) to allow for a more implicit measure of the participant attitudes to verify their preferred variety. Here, the participants were asked to indicate their degree of agreement or disagreement on a 5-level Likert scale. The measures ranged from a strong preference for non-rhoticity (score of 1) to a definite partiality for rhoticity (score of 5), with a neutral opinion falling on 3. As a result, we assumed that a final score higher than 30 would show an American rhotic English preference, while lower than 30 displays an inclination for a British non-rhotic accent. All the information provided in the present study was self-reported by the informants. Since the questionnaire contained questions revealing the purposes of the study, such as speaker’s target norm and attitude toward rhoticity, which may bias the participants’ accentual behavior before the recording sessions, this task was administered as the last step in the experiment.

4.3. Data coding

A total of 762 tokens for 13 speakers were elicited in the two English tasks of the experiment. The tokens did not include the cases that potentially give rise to linking /r/, where postvocalic (r) ordinarily surfaces due to the following vowel-initial word in most non-rhotic varieties. A binary choice was made between the presence and absence of the (r) variant for each token through perceptual inspection by the first author, who is phonetically trained. An additional phonetically-trained listener checked a subsample of the tokens ($n = 200$), which returned a high degree of inter-rater reliability (94% agreement).

Alongside the presence or absence of (r), each test token was also coded for the independent variables that have been examined in studies of rhoticity elsewhere and shown to be influential in order to facilitate the comparability of the present research with others. Furthermore, extra-linguistic factors that were expected to affect the production of (r) in Chinese English were also grouped and analyzed. Below, we will unpack each.

4.3.1. Social factor groups

Speakers of the present study came from 8 different provinces of China, which represents a broad spectrum. Apart from *place of origin*, we identified 9 other potential demographic and speaker-specific variables: *age*, *gender*, *language attitude*, *target norm*, *L1 Mandarin rhoticity*, *speech style*, *media exposure*, *additional language(s) spoken by the informants*, and *experience abroad* and tested whether they determine postvocalic /r/ presence in Chinese English. Table 3 summarizes the distribution of these extra-linguistic variables in our test group.

Factor Group	Factors/Criteria	Number of Participants
Age	Older group (> 45 years old)	5/13
	Younger group (\leq 44 years old)	8/13
Gender	Male	5/13
	Female	8/13
Language Attitude	American English (score in the attitude survey \geq 30)	7/13
	British English (score in the attitude survey < 30)	6/13
Target Norm	American English	7/13
	British English	6/13
Rhoticity in Mandarin	Nonrhotic (> 80% non-rhotic in Mandarin reading task)	8/13
	Rhotic	5/13
Place of Origin in China	North	10/13
	South	3/13
Speech Style	Reading	all speakers
	Free Speech	all speakers
Regular English Media Exposure	American English	3/13
	British English	2/13
	Both	8/13
Languages Spoken	ie. Russian, German, Japanese, etc.	see Table 1
Experience	ie. New Zealand, Australia, US, etc.	see Table 1

Table 3: Summary of external factors

It should be noted that since age and gender differences in rhoticity may be more directly related to the sociohistorical sphere surrounding ESL contexts and the sociolinguistic identities and ideologies therein, these factors are more likely to play a role in Outer Circle Englishes (e.g., Chand, 2010 for Indian English, Hartmann and Zerbian 2010 for Black South African English) than in Expanding Circle Englishes.⁵ However, we expected *speech style* to be a more conspicuous predictor of /r/ presence in Chinese English especially given that both studies on Korean English (Kang, 2013) and Yunnan English (Sundkvist and Gao, 2016) have shown that different speech contexts would trigger various levels of attention, thus resulting in variation in rhoticity in Expanding Circle Englishes. Furthermore, as mentioned above, the variable nature of rhoticity depending on speech style in rhotic varieties of Mandarin Chinese motivated us to include this factor as a predictor. The remaining factors, *language attitude*, *native language rhoticity*, *media exposure*, *additional languages spoken*, and *experience abroad* potentially account for variation in the realization of (r) according to Melchers and Shaw (2003).

4.3.2. Linguistic factor groups

Apart from above-mentioned sociolinguistic factors, phonological constraints have been considered to be influential to (r) as well (e.g. Irwin and Nagy, 2007; Piercy, 2012). Here we focus on 3 phonological variables that are found to be the most robust predictors in various research on rhoticity: *preceding vowel*, *stress* and *syllable structure*. Table 4 lists the factor groups and individual subfactors with an example for each.

Factor Group	Factors	Example
Preceding Vowel	NURSE	<i>emergency</i>
	LETTER	<i>pleasure</i>
	NEAR	<i>nearby</i>
	SQUARE	<i>scared</i>
	FORCE	<i>before</i>
	START	<i>alarm</i>
Syllabic Stress	Stressed	<i>version</i>
	Unstressed	<i>southern</i>
Syllable Structure	Coda	<i>perfect</i>
	Consonant Clusters	<i>shepherd</i>

Table 4: Summary of internal constraints

In terms of the preceding vowel, each test token was coded adopting the citation form of Wells' lexical sets (Wells, 1982). To ensure comparability with previous studies on rhoticity in Inner Circle varieties (e.g., Nagy and Irwin, 2010; Piercy, 2012) as well as Expanding Circle Englishes (e.g., Kang, 2013), we identified 6 preceding vocalic environments, in particular the vowels in NEAR, SQUARE, LETTER, NURSE, START, and FORCE lexical sets.

With regard to the factor *stress*, the variant (r) in the elicited tokens was coded as occurring either in stressed or unstressed syllables. The coding of stress was based on the actual realization of stress in each test token rather than simply following the citation form of the preceding vowels. For instance, the function words *for* and *or* were coded as unstressed, even though their citation form, which contains the FORCE vowel, could lead one to categorize the vowel as stressed. The significant effect of stress has been reported in a number of studies on rhoticity in Inner Circle English varieties, such as Irwin and Nagy's (2007) research on Boston English and Piercy's (2012) study on Dorset English. Both of the studies have indicated that the stressed syllables favor rhoticity while the unstressed syllables disfavor it. Furthermore, Bonnici's (2010) study on Maltese English, an Outer Circle variety, also examined the influence of syllabic stress on (r) presence and showed a similar trend.

The last phonological factor tested was *syllable structure*. The tokens were coded based on the position of postvocalic /r/: (i) the coda position, which is followed by a heterosyllabic consonant as in *nearly*, (ii) word final position before a pause or a consonant-initial word (e.g., *before*), and (iii) before a tautosyllabic consonant, occurring either in the same morpheme as the rhotic sound (e.g., *southern*) or was separated from it

by a morpheme boundary (e.g., *villagers*). It should be noted that syllable structure has also been shown to be an important linguistic factor conditioning (r). In Labov's (1972) department store study, for example, it has been demonstrated that speakers pronounced word-final (r) in *floor* more frequently than the (r) in consonant clusters such as *fourth*. Nagy and Irwin (2010) and Piercy (2012) also showed that (r) in the coda position was produced at a higher rate than (r) in consonant clusters.

5. Results

5.1. Overall degree of rhoticity

Table 5 shows the frequency of (r) presence in all cases, suggesting that Chinese English represented by the English teachers in our study is only slightly rhotic at best (55.1%). This pattern goes slightly in the direction of Kang's (2013) findings on Korean English, although Korean speakers in that study were shown to resort to rhotic realizations with a much higher frequency (73.9%) than in the present study.

	r-full	r-less	Total
N	420	342	762
%	55.1	44.9	100

Table 5: Overall distribution of (r) in Chinese English

Our finding however differs from what was previously shown for Yunnan English, which is based in the South of China. As noted above, Sundkvist and Gao (2016) focused primarily on variation in speech styles concerning this variety. Instead of presenting an overall degree of (r) presence, they reported the percentage of (r) across three different speech types—interview (38.5%), reading (45.1%), and questionnaire (64.9%). Since the present research investigated two similar speech styles that were included in the Yunnan English study, namely interview and reading, we can conclude that Yunnan English is rather non-rhotic compared with Chinese English, with the degree of rhoticity for both speech tasks lower than 50%. A reasonable explanation for this distinction concerns the substrate language Mandarin, in particular the dialectal variation therein. As has been mentioned before, the Mandarin rhotic accent has been widely adopted in North China where a great majority of our participants originated from while Mandarin speakers from the South rarely produce the rhotic suffix since this feature is largely missing in the Southern Chinese dialects, which includes the Yunnan province. As such, the comparison between Yunnan English in Sundkvist and Gao (2016) and (Northern) Chinese English in our study makes a significant contribution to the role of dialectal variation in this domain although a direct comparison would be necessary to reveal the exact magnitude of the difference.

5.2. Inter- and intra-speaker variability

Inter- and intra- speaker variation of rhoticity in Chinese English is shown in Figure 1. As can be seen, the individual proportion of rhoticity varies across participants. More specifically, no Chinese English speaker was found to be either categorically rhotic or non-rhotic. The most rhotic speaker pronounced 98% of tokens as /r/-full, while the most /r/-less participant produced this variant as /r/-full only 27% of the time. The proportion of rhoticity for most participants ranges between 40 to 60%. This pattern accords well with Sandkvist and Gao's findings (2015) on Yunnan English, which reports that six out of eight participants displayed rhoticity to "a variable although considerable degree" (p. 14), while the remaining two were considered mostly non-rhotic.

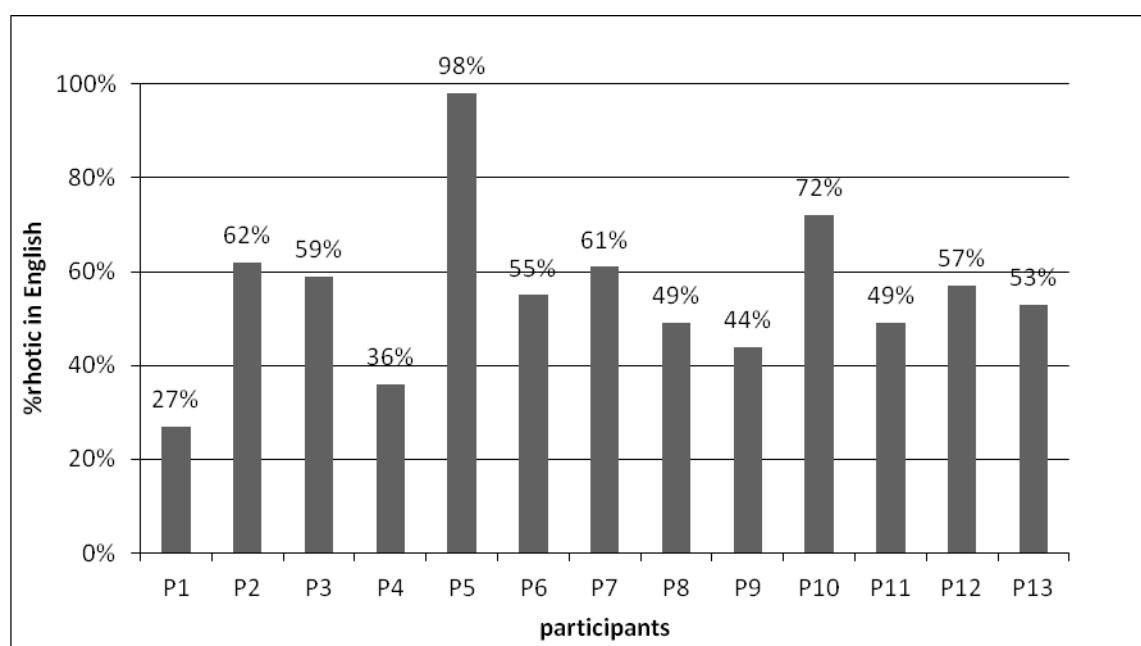


Figure 1: Individual distribution of (r) by Chinese English speakers; P stands for participant.

5.3. Multivariate analysis

Table 7 presents the results of the multivariate analysis by a mixed effects model (best run, stepping-down) using the *Rbrul* package (Johnson, 2009) in R (R Core Team, 2016), with the presence of rhoticity as the application value. An alpha level of .05 was used for all statistical tests. Non-significant factor groups were not presented in the table. In particular, *preceding vowel* was the most significant phonological factor ($p < 0.001$) conditioning rhoticity in Chinese English while both *stress*⁶ and *syllable structure* were not significant. More specifically, an inspection of the factor weights reveal that a preceding NEAR vowel was by far the strongest predictor of (r) presence. This was

followed by the SQUARE, LETTER, and NURSE contexts. The START and FORCE vowels were however disfavored rhoticity.

In terms of external factors, *target norm* turned out to be the most significant variable affecting rhoticity ($p = 0.0018$). Participants with American English as target norm strongly favored rhoticity while those with British English preference disfavored it. Furthermore, variation in speech styles was also shown to exert a significant albeit weak influence on rhoticity ($p = 0.029$). More specifically, the free speech task favored rhoticity while the passage reading task slightly disfavored it. Finally, we found a significant effect for the factor *L1 rhoticity* ($p = 0.0023$). This variable was operationalized by the degree of /r/-fullness that each participant showed in the Mandarin Chinese read aloud task, which was additionally verified by their place of origin. However, the directionality of the influence was contrary to our expectation. In particular, we predicted a positive correlation between the speakers' degree of rhoticity in L1 Mandarin and their L2 English. Instead, the participants in the rhotic Mandarin group strongly *disfavored* rhoticity in English. We will revisit this unexpected finding in our discussion below.

<u>Best run, stepping down</u>				
Dependent variable	N = 762			
= Presence of /r/	Grand mean = 0.551			
	Intercept = 1.456			
	Degree of freedom = 11			
Preceding Vowel (P < .0001)	Log odds	Factor weight	N	% rhotic
NEAR	1.08	.75	91	71.4
SQUARE	0.66	.66	92	70.7
LETTER	0.63	.65	183	69.4
NURSE	0.57	.64	116	70.7
START	-0.97	.28	104	33.7
FORCE	-1.97	.12	176	26.1
Target Norm (P = .0018)	Log odds	Factor weight	N	% rhotic
American English	1.20	.73	428	61.2
British English	-1.20	.27	334	47.3
Task (P = .029)	Log odds	Factor weight	N	% rhotic
Free speech	0.33	.59	294	55.4
Text	-0.33	.41	468	54.9
/r/ in Mandarin (P = .0023)	Log odds	Factor weight	N	% rhotic
nonrhotic	1.00	.73	446	62.3
rhotic	-1.00	.27	316	44.9

Table 6: Significant internal and external constraints on rhoticity in Chinese English

6. Discussion

In this section, we discuss 4 of our variables that turned out to be significant factors in the present study in light of the findings found in previous studies on the variant (r).

6.1. Preceding vowel

Among 3 phonological factors explored in the present study, the preceding vowel context was shown to be the only one that displays a statistically significant effect on rhoticity in Chinese English. The factor weights in Table 7 reveal the following hierarchy in terms of the strength of the vowel predicting /r/-fullness.

- (1) Vowel hierarchy for the variant (r) in Chinese English:
 NEAR > SQUARE > LETTER > NURSE > START > FORCE

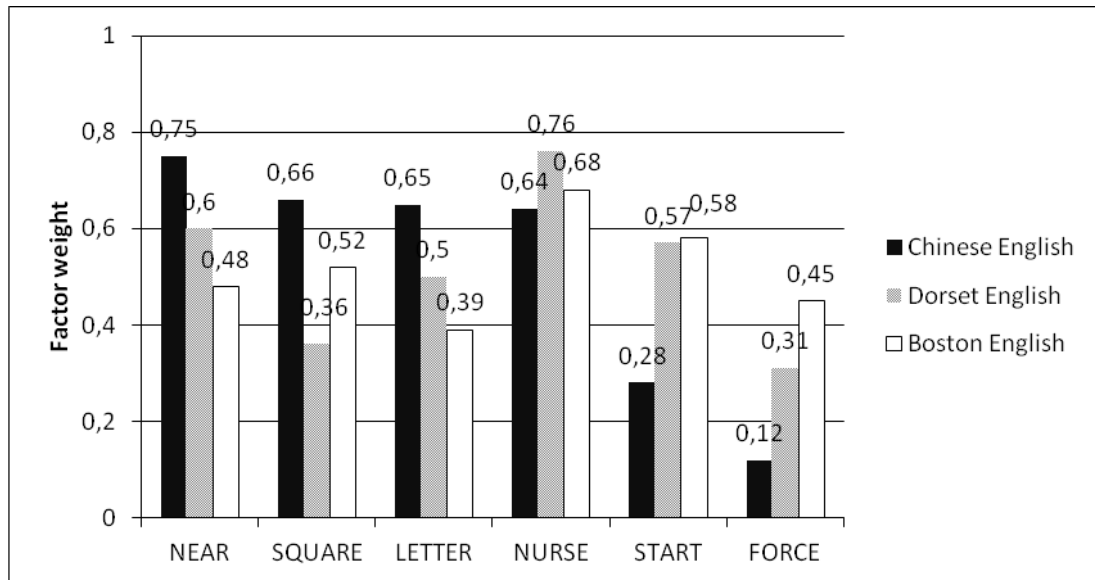


Figure 2: Comparison of the effect of preceding vowel on (r) presence in Chinese English and two Inner Circle varieties

Figure 2 displays a comparison of the effect of preceding vowel on (r) in Chinese English and the two Inner Circle varieties previously studied by Nagy and Irwin (2010) and Piercy (2012). Some commonalities can be observed concerning the effect of the preceding NEAR, NURSE, SQUARE, and FORCE vowels. While a preceding NEAR vowel favors rhoticity in both Chinese English and Dorset English, a preceding SQUARE vowel favors rhoticity in both Chinese English and Boston English. Furthermore, all three varieties favor a preceding NURSE vowel, which may indicate that the stressed schwa favors /r/-fullness in both Expanding Circle and Inner Circle Englishes. According to Nagy and Irwin (2010: 267), the NURSE vowel favors rhoticity more than the unstressed LETTER vowel in almost all studies about American English varieties. This finding, however, is in contrast to the present findings. As the figure shows, Chinese English speakers also favor rhoticity with a preceding LETTER vowel, while Boston English disfavoring it. No clear trend can however be seen for Dorset English. As Irwin and Nagy (2007: 141) notes, the divergent behavior of (r) by preceding environments in different varieties can be due to differences in vowel qualities across these varieties or due to differential coding of vowels in these studies. While this may be true, we suggest that the preference for rhotic realizations after the unstressed schwa in Chinese English finds a more straightforward explanation if we take into account perceptual similarity concerning the morphophonological properties of the L1 and those of L2. In particular, the above mentioned rhotic suffix in Mandarin bears a stark resemblance to English schwa /r/ sequences in that in both languages (i) these sequences occur in unstressed positions, (ii) they are mainly found at the end of a word, and (iii) they mostly function as a suffix (e.g.,

teacher, warmer, etc.). Since Mandarin Chinese rhoticity is primarily suffixal (i.e., it arises when the rhotic suffix is realized), we speculate that /r/-coloring in English in the interlanguage of our participants has found a strong affinity to the rhotic suffix in their L1, thereby yielding a strong preference to render schwa /ɪ/ as overly more rhotic than expected.

Another notable distinction between Chinese English and other two Inner Circle varieties concerns the preceding START vowel. Both Boston English and Dorset English favor (r) in this context while Chinese English disfavors it. The last noteworthy difference involves the preceding FORCE vowel. Although it is shown to be disfavored also in both Boston English and Dorset English, Chinese English displays a stronger dispreference than the two varieties. While we find no direct explanation for the distinction so far, the fact that both START and NORTH are dispreferred is on a par with Feagin's (1990: 137) findings in Anniston, Alabama, where back vowels were shown to disfavor [r] in comparison to front vowels.

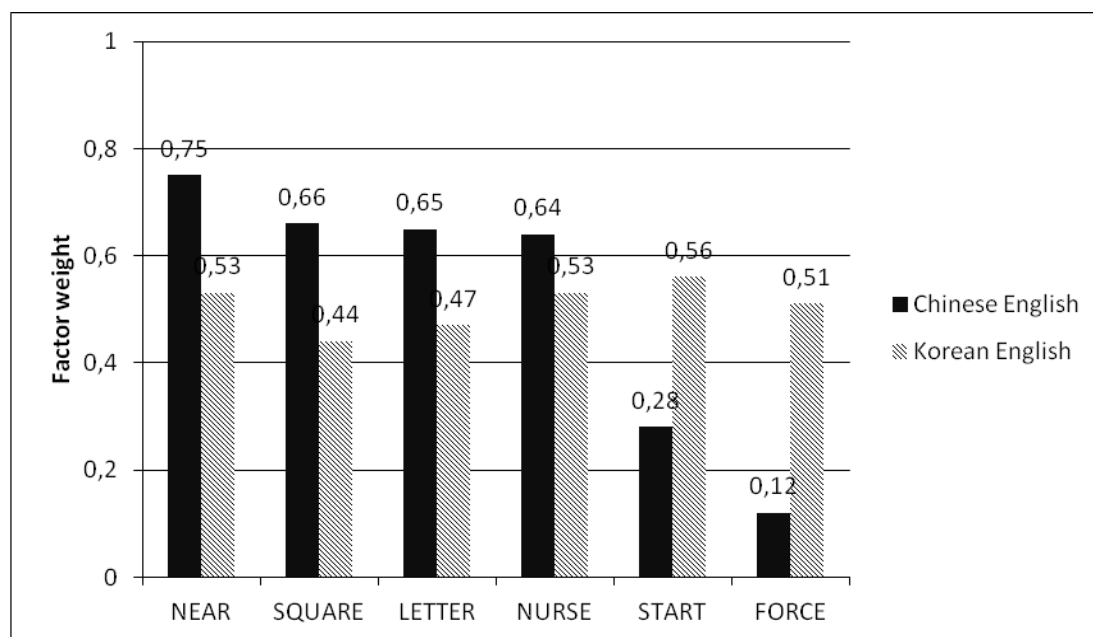


Figure 3: Comparison of the effect of preceding vowel on /r/ presence in Chinese English and Korean English

How do our findings compare to other Expanding Circle Englishes? Figure 3 displays the comparison of the factor weight of preceding vowel on rhoticity in Korean English (Kang 2013) and the present study. First of all, there is no distinct similarity between these two data sets, except for the preceding NEAR and NURSE vowels. In particular, both Korean English and Chinese English favor rhoticity with preceding NEAR and NURSE vowels. A preceding NEAR vowel strongly favors rhoticity in the present study while only slightly favoring rhoticity in Korean English. After the SQUARE and LETTER vowels, while Chinese English favors rhoticity, Korean English disfavors it in the same contexts. By contrast, in Chinese English, both START and FORCE strongly

disfavor rhoticity while they slightly favor rhoticity in Korean English. Since the factor weight of all preceding vowels remains rather neutral in the range of 0.44-0.56 in Kang's results, it is however difficult to detect any robust vowel hierarchy in Korean English unlike what we found for Chinese English.

To conclude, the pattern observed in the present study bears more resemblance to Inner Circle English varieties rather than Korean English, which is also an EFL variety that can be assumed to form another important Expanding Circle variety in East Asia. One possible explanation for this difference in the two EFL varieties could be due to the stark differences in the proficiency levels of the speakers tested. Kang's study investigated Korean college students, who we assume to be less proficient and less stable in their English pronunciation than the participants recruited for the present study, who were English language professionals, more specifically college English teachers who use English regularly for teaching and communication purposes. To this end, the proficiency level of informants in the present study is closer to that of native English speakers in comparison to the Korean English learners in Kang's study. Accordingly, the patterning of the vowel preceding (r) in the present study is more similar to the one found in Inner Circle varieties compared with the Korean English variety.

6.2. Target norm

It should be remembered that the data concerning the participants' self-proclaimed target norm was collected based on the questionnaire. To further confirm whether their preference was consistent, an additional survey of language attitude on their preference for RP non-rhotic accent or American English rhotic accent was conducted via a 5-level Likert scale. Results show that our participants' language attitude and claimed target norms are completely consistent. As expected, target norm showed a significant effect on their rhoticity. Participants with American English as target norm produced more (r) tokens than those targeting British English. The present study is the first of its kind to investigate speakers' target norm as an influential factor determining rhoticity in Expanding Circle Englishes.

6.3. Speech style

Variation in speech styles also plays a significant role in constraining rhoticity in the present study. Results indicate that Chinese English speakers produced more rhotic sounds in free speech than in the reading task. A possible explanation for this result would be the difference in the level of formality between the two tasks. To that end, we assume that the reading task induced a more formal style than the free speech task. As such, our Chinese English speakers can be said to produce more rhotic tokens in the informal style than in the formal one. Indeed, this pattern echoes the production context of rhotic suffixes in their L1 Mandarin. As mentioned before, the rhotic suffix 'ㄌ' in Mandarin Chinese is known to be pronounced more frequently in informal speech than in formal situations. It could be concluded that the stylistic choices concerning the rhotic suffix in

Chinese speakers' L1 Mandarin has a potential impact on the realization of (r) variant in Chinese English.

The direction of stylistic influence on rhoticity in our study is however in contrast to what Sundkvist and Gao (2016: 53-54) reported for Yunnan English. In particular, Yunnan English participants showed the lowest degree of rhoticity (38.5%) in the most informal interview task while producing much higher rates of rhoticity in tasks that triggered more formal styles (45.1% in the reading task and 64.9% in the most formal dialectological-style questionnaire task). The non-rhotic Mandarin spoken by Yunnan English speakers may account for this difference. As discussed above, Yunnan province is located in South China, where the rhotic suffix is rarely used. As a result, we conclude that the productions of Yunnan English speakers could not have been affected by the same stylistic effect as those of the rhotic Mandarin speakers in our study.

6.4. Rhoticity in L1

We predicted that since Mandarin Chinese is more rhotic in North China, Chinese English speakers from such rhotic dialect backgrounds would exhibit a higher proportion of /r/-fullness in English than Mandarin speakers from non-rhotic dialect backgrounds. We employed a Mandarin reading task to evaluate our participants' degree of rhoticity in their L1. We resorted to a categorical distinction in L1 rhoticity (i.e., non-rhotic vs. rhotic) since 8 out of 13 participants did not realize the rhotic suffix in at least 8 of the 10 target Mandarin Chinese words, which we took to be a good indicator of strong non-rhoticity in the L1. Accordingly, we labeled these participants as "non-rhotic". The remainder (n = 5) was automatically categorized into the rhotic Mandarin group. Although L1 rhoticity was a significant factor in the model,⁷ the variable *target norm* exerted a greater influence on the production of (r), as discussed above. We suggest that since college English teachers have years of experience in learning and teaching English, their English speech patterns are less likely to be subject to L1 influence, thereby exhibiting more convergence with the phonological patterns of the respective target norms. We speculate that the substratum language effect may be stronger in the accents of EFL *learner* varieties than more stable types of English found in the Expanding Circle.

7. Conclusion and outlook

Although the speakers of the EFL variety investigated here form only a small portion of a massive body of EFL speakers in East Asia, it is representative of the type of English taught and used at EFL instructional settings in the region. Consequently, the English represented by our college teachers is acquired by an ever-growing learner population in China, which will eventually yield masses of highly proficient users of English in the region. To that end, we explored a relatively small but quite instructive aspect of this emerging variety, the degree of rhoticity, and showed the way internal and external factors constrain the production of (r) on the basis of experimental data collected from college English teachers in China. Our results suggest that Chinese English represented

by these teachers is slightly rhotic and thus sounds closer to American English accent, albeit with considerable inter- and intra- speaker variability. The preceding vowel environment was shown to be the most significant phonological factor conditioning the production of the rhotic variant. Specifically, a preceding NEAR vowel was found to be the strongest predictor of (r) in Chinese English while the preceding FORCE vowel strongly disfavored rhoticity. Upon closer inspection, we showed that the preceding vowel pattern observed to condition rhoticity in our study is strikingly closer to the pattern reported in previous studies on Inner Circle varieties than Expanding Circle Englishes. In and of itself, this is a clear indication that we have dealt with an already established, relatively stable nonnative variety, which is likely to form one of the important sources of input and serve as a salient model of English pronunciation for incoming English learners in China.

What does the systematic emergence of similar structural factors in this nonnative variety, which also modulate the patterning of rhoticity in other English varieties imply for the emergence of novel sound patterns and the way it should be studied? Our study has shown that the variety under investigation here permeates the classic boundaries of the field of Second Language Acquisition, where the nature and sources of *non-convergence* between nonnative and native users dominate various lines of empirical inquiry and consequently shape theoretical approaches. What we have instead observed here is a systematic variability that is on the one hand reminiscent of native varieties of English, and seemingly more stable and complex than that found in SLA studies on the other. This resonates with the Dynamic Systems Theory (DST) approach to multilingualism and second language acquisition, where languages (including second or additional languages) are viewed as complex and dynamic systems, which are argued to be in a constant flux (Herdina and Jessner, 2002; De Bot et al., 2007). For instance, De Bot et al. (2007: 19) suggest that “through iterations of simple procedures that are applied over and over again with the output of the preceding iteration as the input of the next, complexity in language emerges”. Our study provides an instructive case to this view, where the target norm turned out to be the most significant sociolinguistic constraint determining (r), crucially overriding cross-linguistic influence effects. In particular, although the patterning of rhoticity in Chinese English was modulated to some extent by the L1 of the speaker, the overall degree of rhoticity for these highly proficient English speakers was more strongly determined by the target norm. This outcome accords with previous research on second language phonological grammars, which has time and again suggested that such nonnative phonological grammars are better understood in conjunction with natural laws of phonetics, perceptual and articulatory constraints on speech processing, as well as a range of cognitive, psychological and social phenomena such as attitudes, motivation and age (see Altmann and Kabak, 2010; Moyer, 2013 for reviews). However, such factors are not solely reserved for nonnative systems since a large array of sociophonetic research has suggested that similar social and psychological factors also govern variability and change within and across lifespans. Accordingly, Kabak (to appear) expands the study of phonological patterns to *all* spheres and facets of language behavior since they are uniformly governed by the same set of phonetic,

cognitive and social constraints although their weights and magnitude of influence may differ from one another. Arguing against the dichotomy between *natural* and *unnatural* patterns of sound change (e.g., Blevins, 2006), Kabak shows that the same external and internal network of variables that operates in diachrony also operates in second language acquisition of phonological systems. To that end, combining diachronic and psycholinguistic evidence, he offers a Dynamic Equational Approach towards sound patterns. The central tenet of the approach is that *all* sound patterns (unless they are highly fossilized) are characterized by a degree of gradience and variability, and that this gradience arises by a complex interplay of both language-internal and language-external factors with phonological variants in all facets and shapes of language evolution and change. Here, the term *equation* encapsulates both an empirical dimension as well as metaphorical extension. That is, it reflects not only the complex relations among the many variables involved in the generation of a sound pattern, but also pushes the idea that sound patterns are best explored when all other variables that co-determine the shape, direction, and magnitude of sound structure are treated as equally unknown as in a mathematical equation. Essentially, the assigned values to the variables within an equation lead to a solution. Furthermore, “equation” resonates equality in treating all forms of language behavior as uniformly *real* such that phonological variation and change due to language contact (be it realized in loanword phonology, first language attrition, and the emergence of interlanguages or creoles, etc.) should be viewed as natural. Accordingly, different forms of language contact are essentially different reincarnations of a dynamic interaction of linguistic systems that are expected to interact with one another, each being subject to the same set of laws. The current study provides a compelling addition to the Dynamic Equational Model in that the same set of linguistic and extra-linguistic constraints that underlie the patterning of rhoticity in both Inner Circle and Outer Circle Englishes (e.g., vowel context, target norm) and in the substrate language (e.g., dialectal rhotic-nonrhotic variability, stylistic variation) have also been observed to be at work in an Expanding Circle variety, albeit with a different ranking and magnitude that are commensurate with the “noise” that the complex interaction factors that SLA and language contact situations bring about.

Before we conclude, a few words are in order concerning the weaknesses of our study that we hope to guide future research. We used only a reading task to confirm the levels of rhoticity in the L1 of our participants. Provided that rhoticity in Mandarin is sensitive to the variation in the formality of speech styles, a simple formal reading task may have resulted in an overall decline of the presence of /r/-full tokens in Mandarin. Future studies should thus employ a variety of speech elicitation tasks that tap different styles of speech in order to reach a more comprehensive score of participants’ rhoticity levels in their L1. One of the empirical contributions of our study, despite this weakness, is that inherent variability within the L1 may mirror in the extent of variability in the L2. That is, transfer due to cross-linguistic influence does not always take place in a whole-sale manner but rather it may reflect the gradient influence of constraints already existing within the L1.

Another problem concerns the test population we studied. In particular, we tested college English teachers in China who were considered to be highly proficient and

representative Chinese English speakers. In apparent time, Chinese English will grow by the number of highly proficient users of English, who are currently EFL students (of these teachers). Thus, future research should include a direct comparison between Chinese EFL learners of English and the norm providers, i.e., English language professionals and teachers, as this may contribute a broader understanding of how accentual norms perpetuate and how phonological input turns into intake by learners.

All in all, the significance of the present paper stems from the fact that it provides a comprehensive empirical study of rhoticity in Chinese English and extends the inquiry into *both* linguistic and extra-linguistic constraints. It contributes to our understanding of the conditions that modulate (r) occurrence in Expanding Circle Englishes, a research area that is severely understudied in comparison to English varieties in the Inner and the Outer Circle. In addition, as rhoticity is considered one of the most notable variables classifying world Englishes, the phonological findings also provide a valuable piece to the burgeoning puzzle of the typology of World Englishes. Since the research on Chinese English phonology is still in its infancy, we expect our study to draw more attention to the scarcity of empirical work on the linguistic and sociolinguistic aspects of this thriving variety. Finally, we hope that our study hinted at a necessary paradigm shift in our current understanding of and approach to World Englishes such that the field should be more directly informed by psycholinguistic and sociolinguistic theories, approaches and insights from the field of Second Language Acquisition. Thus, we suggest that an “L2 acquisitionist turn” is inevitable in the conceptualization of alternative Englishes.

Notes

1. Whenever necessary, we will use /r/ to mark the rhotic phonemic category in English although there is considerable variation in its exact phonetic realization across English varieties as well as in different phonological contexts within individual varieties (see Gimson and Cruttenden, 1994: 186-190).

2. Following the sociolinguistic research conventions, we interchangeably refer to the variable realization of postvocalic /r/ as the “variant (r)” throughout the paper.

3. English spoken or written by people from mainland China has not yet found a consistently used term in the field. Citing various sources, He and Li (2009: 71) list ‘Chinglish’, ‘Chinese English’, ‘Sinicized English’, as well as ‘China English’ as labels previously used in the literature. Although all of the participants in the present study had been living in North China at the time of testing, we will use ‘Chinese English’ to refer to their variety of English since they originated from various places in China.

4. Indeed, language professionals who teach and test English language skills in EFL contexts can be said to create “codification” practices which serve as the standard for English education. In this vein, Shim (1999: 250-256), for example, argues that what is learned in Korean schools is different from the target norm, American English, and observes that Korean EFL users in the country now share a set of unique lexico-semantic, pragmatic and grammatical rules, which feature in English textbooks used in Korean high schools.

5. We chose 44/45 as the boundary to reflect the mean of the age range (31-59) we had in our population. While this cut allowed us to have two groups with a relatively balanced number of participants in each, it also constitutes a sensible division from the point of view of major

technological, social and political changes that were incipient in the 1990s, an era characterized by, among others, the rise of multiculturalism, social media and the World Wide Web, which arguably exerted a significant influence on the spread of English around the world (we assumed that those teachers younger than 45 at the time of testing were university students or in the early phases of their professional training in the 1990s and beyond).

6. The fact that stress did not come out significant is perhaps not surprising given that stress was highly co-linear with vowel quality.

7. It should be noted again that the fact that we observed a high rate of rhoticity with the LETTER vowel in our sample can be attributed to a potential cross-linguistic influence from Mandarin Chinese, where the rhotic suffix is phonetically similar to English schwa /r/.

Acknowledgements

The authors would like to thank the teachers who participated in this study at the North China Institute of Science and Technology, as well as Marie-Christin Himmel at the University of Würzburg and the two anonymous reviewers for their helpful comments.

References

- Altmann, Heidi and Barış Kabak (2010): "Second language phonology". In N. Kula et al., eds., *Continuum Companion to Phonology*. London: Continuum, 298-319.
- Becker, Kara (2009): "/r/ and the construction of place identity on New York City's Lower East Side". *Journal of Sociolinguistics*, 13(5): 634-658.
- Becker, Kara (2014): "(r) we there yet? The change to rhoticity in New York City English". *Language Variation and Change*, 26(2): 141-168.
- Blevins, Juliette (2006): "New perspectives on English sound patterns: 'Natural' and 'Unnatural' in Evolutionary Phonology". *Journal of English Linguistics*, 34: 6-25.
- Bonnici, Lisa Marie (2010): *Variation in Maltese English: The Interplay of the Local and the Global in an Emerging Postcolonial Variety*. University of California, Davis.
- Chand, Vineeta (2010): "Postvocalic (r) in urban Indian English". *English World-Wide*, 31(1): 1-39.
- Chen, Ping (1999): *Modern Chinese: History and Sociolinguistics*. Cambridge University Press.
- Gimson, Alfred Charles and Alan Cruttenden (1994): *Gimson's Pronunciation of English*. 5th edition. London-New York-Sydney-Auckland: Edward Arnold.
- De Bot, Kees, Wander Lowie and Marjolijn Verspoor (2007): "A dynamic systems theory approach to second language acquisition". *Bilingualism: Language and Cognition*, 10(1): 7-21.
- Deterding, David (2006): "The North Wind versus a Wolf: Short texts for the description and measurement of English pronunciation". *Journal of the International Phonetic Association*, 36(2): 187-196.
- Deterding, David (2007): *Singapore English*. Edinburgh University Press.
- Duanmu, San (2007): *The Phonology of Standard Chinese*. Oxford University Press.
- Eberhardt, Maeve and Corinne Downs (2015): "(r) You Saying Yes to the Dress? Rhoticity on a Bridal Reality Television Show". *Journal of English Linguistics*, 43(2): 118-142.

- Feagin, Crawford (1990): "The dynamics of a sound change in Southern States English: From R-less to R-ful in three generations". In Jerold A. Edmonson, Crawford Feagin, and Peter Mühlhäusler, eds., *Development and Diversity: Linguistic Variation across Time and Space*. Dallas: Summer Institute of Linguistics/ University of Texas at Arlington, 29-146.
- Fowler, Joy (1986): *The Social Stratification of (r) in New York City department stores, 24 years after Labov*. Unpublished Ms, pp. 268.
- Giegerich, Heinz J. (1992): *English Phonology: An Introduction*. Cambridge: Cambridge University Press.
- Gut, Ulrike B. (2004): "Nigerian English: Phonology". In E.W. Schneider et al., eds., *A Handbook of Varieties of English. Volume 1: Phonology*. Berlin: Mouton de Gruyter, 813-830.
- Hansen, Jette G. (2001): "Linguistic constraints on the acquisition of English syllable codas by native speakers of Mandarin Chinese". *Applied Linguistics*, 22(3): 338-365.
- Hartmann, Dieter and Sabine Zerbian (2010): "Rhoticity in Black South African English—a sociolinguistic study". *Southern African Linguistics and Applied Language Studies*, 27(2): 135-148.
- Hay, Jennifer and Alhana Clendon (2012): "(Non)-rhoticity: Lessons from New Zealand English". In T. Nevalainen and E.C. Traugott, eds., *The Oxford Handbook of the History of English*. Oxford: Oxford University Press, 761-772.
- He, Deyuan and David Li (2009): "Language attitudes and linguistic features in the 'China English' debate". *World Englishes*, 28(1): 70-89.
- Herdina, Philip and Ulrike Jessner (2002): *A Dynamic Model of Multilingualism: Perspectives of Change in Psycholinguistics*, vol. 121. Clevedon: Multilingual Matters.
- Irwin, Patricia and Naomi Nagy (2007): "Bostonians /r/ speaking: A quantitative look at (R) in Boston". *University of Pennsylvania Working Papers in Linguistics*, 13(2): 135-147.
- Johnson, Daniel Ezra (2009): "Getting off the GoldVarb standard: Introducing Rbrul for mixed-effects variable rule analysis". *Language and Linguistics Compass*, 3(1): 359-383.
- Kabak, Barış (forthcoming): "A dynamic equational approach to sound patterns: The (un)stability of English dental fricatives illustrated". In B. Busse and R. Möhlig-Falke, eds., *Patterns in Language and Linguistics*. Mouton de Gruyter (upcoming).
- Kachru, Braj B, ed. (1982): *The Other Tongue: English Across Cultures*. Urbana, IL: University of Illinois Press.
- Kachru, Braj B. (1985): "Standards, codification and sociolinguistic realism: The English language in the Outer Circle". In Randolph Quirk and H. G. Widdowson, eds., *English in the World: Teaching and Learning the Language and Literatures*. Cambridge: Cambridge University Press, 11-30.
- Kang Hyeon-seok (2013): "Internal and external constraints on rhoticity in Korean English". *The Sociolinguistic Journal of Korea*, 21(2): 1-27.
- Labov, William (1966): *The Social Stratification of English in New York City*, Washington, DC: Center for Applied Linguistics.
- Labov, William (1972): *Sociolinguistic Patterns (No. 4)*. Philadelphia: University of Pennsylvania Press.
- Llamzon, Teodoro A. (1997): "The phonology of Philippine English". In M.L.S. Bautista, ed., *English is an Asian Language: The Philippine Context*. Sydney: The Macquarie Library, 41-48.
- Love, Jessica and Abby Walker (2012): "Football versus football: Effect of topic on /r/ realization in American and English sports fans". *Language and Speech*, 56(4): 443-460.

- Mahboob, Ahmar and Nandra H. Ahmar (2004): "Pakistani English: Phonology". *A Handbook of Varieties of English: A Multimedia Reference Tool*, 1003-1015.
- Mather, Patrick-André (2012): "The social stratification of /r/ in New York City: Labov's department store study revisited". *Journal of English Linguistics*, 40(4): 338-356.
- Melchers, Gunnel and Philip Shaw (2003): *World Englishes*. London: Arnold.
- Miller, Corey A. (1998): "R-lessness in Philadelphia". In C.A. Moyer, ed., *Sociolinguistics, Language and Society*. New Delhi, India: Sage, 79-96.
- Moyer, Alene (2013): *The Phenomenon of Non-Native Speech*. Cambridge, UK: Cambridge University Press.
- Nagy, Naomi and Irwin, Patricia (2010): "Boston (r): Neighbo (r) s nea (r) and fa (r)". *Language Variation and Change*, 22(2): 241-278.
- Piercy, Caroline (2012): "A transatlantic cross-dialectal comparison of non-prevocalic /r/". *University of Pennsylvania Working Papers in Linguistics*, 18(2).
- Poedjosoedarmo, Gloria (2000): "A description of the English pronunciation of young educated Singaporeans: A study in multidimensional variation". *The English Language in Singapore: Research on Pronunciation*, 65-75.
- R Core Team (2016): *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org/>
- Salbrina, Sharbawi and David Deterding (2010): "Rhoticity in Brunei English". *English World-Wide*, 31(2): 121-137.
- Schmied, Josef (2006): "East African Englishes". In Braj B. Kachru, Yamuna Kachru and Cecil L. Nelson, eds., *The Handbook of World Englishes*. Oxford: Blackwell, 188-202.
- Schneider, Edgar W. (2003): "The dynamics of New Englishes: From identity construction to dialect birth". *Language*, 79(2): 233-281.
- Schneider, Edgar W. (2007): *Postcolonial English: Varieties around the World*. Cambridge University Press.
- Setter, Jane, Cathy S. P. Wong and Brian H.S. Chan (2010): *Hong Kong English*. Edinburgh: EUP.
- Shim, Rosa Jinyoung (1999): "Codified Korean English: process, characteristics and consequence". *World Englishes*, 18(2): 247-258.
- Sullivan, Anthea E. (1992): *Sound Change in Progress: A Study of Phonological Change and Lexical Diffusion, with Reference to Glottalization and r-loss in the Speech of Some Exeter Schoolchildren*, vol. 17. Exeter: University of Exeter Press.
- Sundkvist, Peter and Man Gao (2016): "Rhoticity in Yunnan English". *World Englishes*, 35(1): 42-59.
- Tan, Chor Hiang and Anthea Fraser Gupta (1992): "Post-vocalic /r/ in Singapore English". *York Papers in Linguistics*, 16: 139-152.
- Tan, Ying Ying (2012): "To r or not to r: Social Correlates of /r/ in Singapore English". *International Journal of the Sociology of Language*, 218: 1-24.
- Thomas, Erik R. (2007): "Phonological and phonetic characteristics of African American vernacular English". *Language and Linguistics Compass*, 1(5): 450-475.
- Trudgill, Peter and Jean Hannah (1994): *International English: A Guide to Varieties of Standard English*. London: Arnold.
- Villard, Sarah (2009): "Postvocalic /r/ in the Upper Valley of Vermont and New Hampshire". Paper presented at NWA 38, Ottawa.
- Wells, John C. (1982): *Accents of English*, vol. 1. Cambridge: Cambridge University Press.

Appendix 1 Questionnaire

College and university English teacher questionnaire

We would like to ask you to help us by answering the following questions concerning foreign language learning and teaching. This is not a test so there are no “right” or “wrong” answers. We are interested in your personal opinion. Please give your answers sincerely as only this will guarantee the success of the investigation. Thank you very much for your help!

I. Background information

1. Your gender? male / female
2. How old are you (in years)?
3. How old were you when you started learning English?
4. How old were you when you started teaching English?
5. Where did you grow up? Where is your hometown? (province, city, town)
6. What Chinese dialect(s) do you speak besides Mandarin?
7. What foreign language(s) can you speak besides English?
8. Do you have any experience abroad? If yes, please indicate the country and length of residence.
9. Which kind of Standard English do you prefer to speak and teach, British English or American English?
10. Which kind of media do you regularly have exposure to?
 - American English media (American TV and films, VOA, etc.)
 - British English media (BBC News, British TV and films, etc.)
 - Both
 - Others (CCTV_news, TED, etc.),

II. Now there are going to be statements some people agree with and some people don't. I would like to know to what extent they describe your own feelings of situation. After each statement you'll find five boxes, please put a '✓' in the box which best expresses how true the statement is about your feelings or situation.

	Absolutely true	Mostly true	Partly true partly untrue	Not really true	Not true at all
1. I like rhotic accent in American English very much.					
2. I like r-less accent in British English very much.					
3. British English r-less accent conveys a sense of arrogance.					
4. American English rhotic accent conveys a sense of arrogance.					
5. American English rhotic accent is a sign of prestige.					
6. British English r-less accent is a sign of prestige.					
7. I personally prefer to teach in American English instead of British English.					
8. It seems that American English is more fashionable than its counterpart in Britain.					
9. I think American English is more popular than British English among my students.					
10. In my opinion, American English is more 'standard' than British English.					

THANK YOU VERY MUCH -- WE REALLY APPRECIATE YOUR HELP!

University of Alicante
LexEsp Research Group
(Research Group in ESP Lexicology and
Lexicography and Vocabulary Teaching)

<http://web.ua.es/en/lexesp/main-page.html>

ESPAÑOL VALENCIÀ ENGLISH

 Universitat d'Alacant
Universidad de Alicante

LEXESP

Seminar |

- Start
- **Researchers**
- Research Topics
- Research Offer
- Publications
- Joint publications
- Lecture Series on Lexis 2014
- III LEXESP Seminar
 - Call for Papers
 - Programme
- II LEXESP Seminar
 - Call for Papers
 - Contact
- I LEXESP Seminar
- Contact

Researchers

- BALTEIRO FERNANDEZ, MARIA ISABEL (Head Researcher)
- AGULLÓ BENITO, INMACULADA
- ALCARAZ ARIZA, MARIA ANGELES
- ALESON CARBONELL, MARIA AMPARO
- BELDA MEDINA, JOSE RAMON
- CAMPOS PARDILLOS, MIGUEL ÁNGEL
- CALVO FERRER, JOSE RAMÓN
- MARTINEZ MOTOS, RAQUEL

Language contact in Gibraltar English: A pilot study with ICE-GBR*

Lucía Loureiro-Porto

lucia.loureiro@uib.es

Cristina Suárez-Gómez

cristina.suarez@uib.es

Universitat de les Illes Balears

ABSTRACT

The variety of English used in Gibraltar has been in contact with a number of European languages, such as Spanish, Italian, Hebrew and Arabic (Moyer, 1998: 216; Suárez-Gómez, 2012: 1746), for more than 300 years. Studies of this variety have traditionally been based on interviews and observation (e.g. Moyer, 1993, 1998; Cal Varela, 1996; Levey, 2008 2015; Weston, 2011, 2013, etc.), and a detailed morphosyntactic description is yet to be published. In this context, the compilation of a reliable Gibraltar corpus using the standards of the *International Corpus of English* (ICE) will constitute a landmark in the analysis of this lesser known variety of English. In the present paper we describe the ICE project and the current state of the compilation of ICE-GBR. In addition, we present a detailed comparison between the section on press news reports of ICE-GB (standard British English) and ICE-GBR, with the aim of identifying morphosyntactic features that reveal the influence of language contact with Spanish in this territory. We explore variables such as the choice of relativizer (assuming a higher preference for *that* in GBR, in agreement with Spanish *que*, the most frequent relativizer, Brucart, 1999: 490), the use of titles and pseudo-titles preceding proper names (which, as shown by Hundt and Kabatek, 2015, are very frequent in English journalese and extremely infrequent in Spanish), and the frequency of the passive voice (expected to be lower in ICE-GBR), among others. A preliminary analysis of these variables reveals that the influence of Spanish on the variety of English used in the Gibraltar press, at the morphosyntactic level, is almost non-existent, limited to occasional cases of code-switching between the two varieties. We hypothesize that a possible

explanation for this strong exonormative allegiance to British English, at least in press news reports, can be found in a strong editorial pressure to reflect the prestigious parent-variety.

Keywords: Gibraltar English, morphosyntactic variation, language contact, International Corpus of English

1. Introduction

Gibraltar lies at the southern tip of the Iberian Peninsula, with a surface of some 6 km² and a population of 32,000 people (Census 2012). It is a British Overseas Territory, and has been subject to the sovereignty of the United Kingdom since the beginning of the 18th century, when Spain, following the War of the Spanish Succession, ceded the territory to Great Britain as part of the Treaty of Utrecht. From that moment, English became the official language, although, due to the Spanish roots of part of the population, as well as for obvious geographical reasons, Spanish remains one of the languages of Gibraltar, spoken with its own distinct Andalusian accent.

Gibraltar has historically been considered “a melting pot of peoples from different cultural backgrounds and with different languages who have settled there throughout the centuries to pursue various military, trade and commercial interests” (Moyer, 1998: 216),¹ which developed a pidgin language for communication. In the 20th century the local community has become ethnically more homogeneous, competent in both English and Spanish, and also speaking *Yanito*, the local vernacular language of Gibraltar, defined as “an Andalusian Spanish-dominant form of oral expression which integrates mainly English lexical and syntactic elements as well as some local vocabulary” (Levey, 2008: 3; see also Moyer, 1998: 216). These days an increase in the use of English at home (inter-parental situation) among youngest Gibraltarians is observed (Kellerman, 2001: 91-93; Levey, 2008: 58, 95-98; Weston, 2013).

The linguistic situation in Gibraltar has been described from different perspectives, such as the use of code-switching (e.g. Kramer, 1986; Moyer, 1993, 1998), its specific phonological features (e.g. Cal Varela, 1996, 2001), the status of Gibraltarian English as regards other Englishes according to Schneider’s (2007) model (e.g. Weston, 2011), as well as in broader and more holistic ways (e.g. Levey, 2008). Nevertheless, all these studies are based on interviews and the observations of researchers, which are always potentially subject to a degree of subjectivity or unintended bias. In order to avoid this, a corpus-based methodology is the preferred option for studies on language variation (see Rissanen (2014) for a good review). The existence of a Gibraltarian corpus, then, would guarantee objectivity in the analysis of this linguistic variety. The team to which the authors belong is currently compiling the Gibraltar component of the *International Corpus of English* (ICE), an international project which seeks to provide corpora which allow for comparisons between different varieties of English.² Although the corpus is still in progress, we present here a preliminary study using one of the completed sections, which includes press news reports; in this context we provide a comparison of the British

and the Gibraltarian press, with the aim of detecting evidence of contact with Spanish. The paper is structured as follows. Section 2 accounts for the language contact situation in Gibraltar. Section 3 describes the ICE project and the current status of ICE-Gibraltar (ICE-GBR). Section 4 describes the data and methodology. Section 5 presents the corpus-based analysis of press news reports in Britain and in Gibraltar. Finally, Section 6 summarizes the results and reaches some conclusions.

2. Language contact in Gibraltar

The linguistic situation in Gibraltar has been the object of research since the 1980s, and has generated an extensive literature (cf. especially Ballantine, 1983, 2000; Cal Varela, 1996, 2001; Cavilla, 1990; Kellerman, 2001; Kramer, 1986; Levey, 2008, 2015; Lipski, 1986; Moyer, 1993, 1998; Weston, 2011, 2013; Suárez-Gómez, 2012). Most studies here have looked at the historical coexistence of English and Spanish and the resulting code-switching variety known as *Yanito*.

As noted in the previous section, the community of Gibraltar has historically been defined as a cultural, military, commercial and linguistic “melting pot” (Moyer, 1998: 216). In fact, official documents report the following populations in the 18th century: “British 434; Genoese 597; Jews 575; Spaniards 185, and Portuguese 25”, showing the coexistence of “peoples from different cultural backgrounds”³ and with different languages. Hence, for communicative purposes it was necessary to develop some sort of pidgin language, in this case a mixture of Italian, Spanish, English, Arabic and Hebrew (Moyer, 1993: 85), which until the end of last century was frequently used by local speakers. By that time, Gibraltar had become a more homogeneous ethnic group, as reflected in the 2012 official census, which reports that 79% of the population is Gibraltarian (Gibraltarians 25,444; UK and Other British 4,249; Moroccan 522 and Other Nationalities, 1979) (<<https://www.gibraltar.gov.gi>>, last access 25 September 2017).

As expected, the historical distribution of the population in Gibraltar correlates with the situation of language use. From a linguistic point of view, Gibraltarians in their 60s and older tend to be relatively competent in both English and Spanish, but currently speak *Yanito*. This is a linguistic variety which identifies Gibraltarians⁴ and which has emerged as a result of “various patterns of code-switching with a proportionally small lexical substratum from Italian, Hebrew and Arabic and a local vernacular” (Moyer, 1998: 216). However, younger generations are more likely to exhibit linguistic code-alternation between *Yanito* and English, most of them with a passive familiarity of Spanish, but communicating mainly in English.⁵

English has been the target language given that, in terms of socioeconomic dominance, it is more prestigious than Spanish. After so many years of contact with different languages, and especially intense between Spanish and English, a nativized variety is acknowledged, labelled Gibraltar(ian) English in the literature (Kellerman, 2001; Levey, 2008, 2015; Weston, 2011, 2015; Seoane et al., 2016). English is becoming not only the official language, but also the first language of most Gibraltarians, and there is explicit recognition of an emerging variety of English which is in the process of

becoming nativized. The English of Gibraltar is turning into a ‘new New English’ (Kellerman, 2001), this also known as Gibraltarian English. Although still absent from New Englishes studies, it is a variety developed to represent a local identity (Le Page and Tabouret-Keller, 1985) and is initially exonormative and with covert prestige, although now gradually turning into a more prestige variety used in formal situations. As a consequence of nativization, we assume that the resulting variety is affected by, and undergoes changes in comparison with, other languages that enter into the contact process (Schneider, 2007). A key element here is the degree of such interaction and the extent to which it has had an effect on the new variety of English.

As a result of contact between coexisting languages, “complex patterns of contact linguistics, including lexical transfer, code switching and code mixing, and discursal and syntactic change and accommodation” (Bolton, 2006: 261) are very likely to occur (see also Thomason and Kaufmann, 1988). Therefore, following Thomason (2001: 63) we assume that in situations of language contact all language levels can be affected and “anything” can be adopted from the languages in contact, from vocabulary, to phonological or structural features. However, it is generally agreed that language contact is especially obvious in the case of lexis and phonology, but less so in syntax and pragmatics, which entails a very intense contact, as shown on Thomason and Kaufmann’s borrowing scale (1988: 74-76). With Gibraltar English, the contact between the two languages has been so intense that the resulting variety has its own peculiar traits not only at the level of lexis, but also in terms of pronunciation and prosody, as observed by Levey (2015: 61), “it has a syllable-timed rhythm rather than a stress-timed one and weak forms are rarely used” (see also Kellerman, 2001: 307-308). Lexically speaking, the resulting variety reflects interference between Spanish and English,⁶ and the semantic fields most affected are cultural terms relating to food (e.g., *greivi* ‘gravy’, *saltipina* ‘salted peanuts’), specialized vocabulary related to docks and constructions (e.g., *cren* ‘crane’, *doquia* ‘dockyard’), vocabulary associated with the classroom (English as the language of education; e.g., *cho* ‘chalk’) and the use of false friends (e.g., *aplicacion* ‘job application’).

3. The *International Corpus of English* project and ICE-Gibraltar

3.1. The ICE project

The *International Corpus of English* (ICE) project was born three decades ago, after a call in 1988 by Sidney Greenbaum (University College London) for English scholars to expand the scope of computerized corpora beyond Brown and the Lancaster-Oslo/Bergen (LOB) corpus, the American and British corpora from the 1960s that had set the standard for corpus-based linguistic studies. In an often quoted note published in *World Englishes*, he stated:

We should now be thinking of extending the scope for computerized comparative studies in three ways: (1) to sample standard varieties from other countries where English is the first

language, for example Canada and Australia; (2) to sample national varieties from countries where English is an official additional language, for example India and Nigeria; and (3) to include spoken and manuscript English as well as printed English. (Greenbaum, 1988: 315)

Since then several ICE components have been launched: ICE-GB (1988), ICE-EA (1999), ICE-IND (2002), ICE-SIN (2002), ICE-PHI (2004) and ICE-HK (2006), among others. Each of them has the same design and consists of 1 million words (60%, spoken material, 40%, written material), distributed in text-types as follows:

SPOKEN (300)	Dialogues (180)	Private (100)
		Public (80)
	Monologues (120)	Unscripted (70)
		Scripted (50)
WRITTEN (200)	Non-printed (50)	Student writing (20)
		Letters (30)
	Printed (150)	Academic writing (40)
		Popular writing (40)
		Reportage (20)
		Instructional writing (20)
		Persuasive writing (10)
		Creative writing (20)

Table 1: ICE internal structure (numbers in brackets indicate number of 2,000-word texts)

Among the most remarkable strengths of the ICE project are that, firstly, 60% of each corpus comprises spoken material (much of it being spontaneous), as seen in Table 1, which allows for the study of less stylized varieties where more international variation is to be expected. Secondly, the project has allowed English corpus linguistics to become a global endeavour, instead of being restricted to the mainstream varieties, British and American English, this expansion being at the heart of Greenbaum's proposals (see also Nelson, 2006: 736-740). It contains parallel corpora of varieties of English as a Native Language (e.g., Canada, New Zealand, Ireland), English as an institutionalized Second Language (i.e., official or widely used language for intranational communication, such as education, media, administration, e.g., India, Singapore English), English as a Second Dialect (ESD, e.g., Jamaican English) and varieties of English spoken in places where its exact status is debatable (e.g., Maltese English). All the corpora are freely available for download following communication with the coordinator.⁷

The size of ICE, one million words per variety, may be considered small by comparison with current standards (see BNC, with 100 million words, COCA, with 520 million words, etc.). Yet it is sufficient for the study of grammatical variation, as shown by Seoane and Suárez-Gómez's (2013) study of the perfect meaning in Asian varieties,

by Suárez-Gómez (2014), which analyses relative clauses, and by Loureiro-Porto (2016), which focuses on modal verbs. However, if the aim is to study low frequency items, such as lexical elements, specific collocations or contractions and even syntactic constructions (e.g., García-Castro 2017), ICE's size is clearly one of its main disadvantages. For such studies, the use of a larger corpus, such as the *Global Web-Based Corpus of English* (GloWbE, Davies, 2013), is required, although important differences regarding the degree of representativeness have been found between ICE and GloWbE (Loureiro-Porto, 2017).

Other caveats regarding ICE include the time lag between corpora (e.g., between ICE-GB and ICE-Nigeria there is a 20-year difference) and even within corpora (e.g., ICE-Fiji contains a 20-year time gap between some of its text types). Such gaps may involve the risk of interpreting diachronic differences as differences between the varieties themselves (cf. Hundt, 2015), especially in the study of rapidly-changing features, such as quotatives (e.g. *be like*) or intensifiers (*super, über*, etc.). In addition, the inter-corpora comparability of registers may be threatened by cultural differences, as explained by Hundt (2015: 384-385; see also Schaub, 2016), who cites the example of a student essay in ICE-PHI consisting of scattered thoughts and incomplete sentences. Spoken private registers also illustrate the effects of cultural differences in the compilation of the corpora, because very often the compiler has no choice but to interview speakers, with the spontaneous and natural character of the text-type then becoming questionable (cf. Hundt, 2015). Additionally, it is not always easy to decide exactly who qualifies as a speaker of the variety being compiled (Mukherjee and Schilk, 2012: 191).

In spite of all these caveats, ICE continues to be the only project that provides representative corpora of varieties of English, although excluding online texts (and important differences have been found between those two ways of compiling corpora, see Loureiro-Porto, 2017). In addition to the 12 corpora released thus far (India, New Zealand, Singapore, Australia, Canada, Great Britain, East Africa, Hong Kong, Ireland, Jamaica, Nigeria and Philippines), the written components of three varieties are also available (Ghana, Sri-Lanka and USA), and 12 international teams are working on the compilation of new members of the ICE family: Bahamas, Fiji, Gibraltar, Malaysia, Malta, Namibia, Pakistan, Puerto Rico, Scotland, South Africa, Trinidad & Tobago and Uganda. With this in mind, ICE-Gibraltar (ICE-GBR) will be described in the following section.

3.2. ICE-Gibraltar: current status

In 2014 the research unit *Variation in English Worldwide* (<<http://view0.webs.uvigo.es>>, Principal Investigator: Elena Seoane) was commissioned to compile the Gibraltar component of the ICE project. The current team comprises three full-time members, three affiliated members, and six PhD students, one of whom also works as a research assistant. During the period 2014-2017 it received funding from the Spanish Ministry of Economy and Competitiveness (grant FFI2014-53930-P) to begin compiling the corpus. In accordance with other ICE corpora, we decided that the written part would be compiled first, and, for practical reasons, the first registers were Reportage (Press News Reports,

W2C in Table 1 above), Persuasive Writing (Press editorials, W2E in Table 1) and Creative Writing (Novels and Short Stories, W2F in Table 1).

Following the ICE structure, 20 texts of 2,000 words each were necessary for Press News Reports, and 10 texts of 2,000 words for Press Editorials. Both of these registers have been extracted from online Gibraltar newspapers such as *Gibraltar Chronicle* and *Panorama*. Nevertheless, the conversion of these electronic texts into part of the corpus requires much more work. To begin with, it is necessary to confirm that the journalists and writers in question are reliable Gibraltar speakers and that they meet the criteria as laid down by ICE:

The authors and speakers of the texts are aged 18 or above, were educated through the medium of English, and were either born in the country in whose corpus they are included, or moved there at an early age and received their education through the medium of English in the country concerned. [The ICE Project 2009]

Since Gibraltar is a very small community (see section 1), it is very common for its young citizens to study abroad, particularly in the UK (as noted by Jennifer Ballantine of the Postgraduate Institute of Gibraltar and Mediterranean Studies and also director of the Garrison Library, Seoane, 2017). Such a context would favour the accommodation of these Gibraltar speakers into British English, as described in the literature on language contact (e.g. Bolton, 2006: 261). Obviously, we need to avoid including texts which do not portray the *real* Gibraltar variety; yet, for the reasons explained by Jennifer Ballantine, it is highly likely that anyone with a university degree has spent some time in the UK. Therefore, in an attempt to find a balance between rigour and pragmatism, we decided that the conditions for anyone to be considered a *true* Gibraltar English speaker should be: (i) they cannot have lived in the UK longer than 10 years, and (ii) they must have been back in Gibraltar for more than 8 consecutive years. These periods would allow citizens to have obtained a university degree in the UK, but at the same time would have been expected to be fully readjusted to Gibraltar speech and language use. These characteristics apply not only to the writers of the texts, but also to any speakers who may have produced utterances included as a quotation in any item of reportage included.

Having identified those news reports and editorials that satisfied the above criteria, the second step in the compilation of sections W2C and W2E involved a random selection of texts which reached 2,000 words, as specified by ICE. Because many news reports are shorter than this, in some cases several texts had to be used as sub-parts of the same file.

The third step involved the xml codification of the texts according to ICE norms. A comprehensive list of codes is provided for compilers, and includes issues such as metadata (e.g. author's age and sex), typographic markup (e.g. bold or italics), content markup (e.g. headings, paragraphs), extra-corpus material (e.g. pieces of text produced by interviewers), and normalization of the text (e.g. annotation of misspellings). This has been done manually with the annotation software Oxygen, as shown in Figure 1. In order to guarantee the correct and coherent annotation of the texts, two individual revisions were made of each transcription. This involved decisions such as the following:

- Distinguishing foreign text from an indigenous use of Spanish forms, as seen in examples (1) and (2) respectively:
 - (1) like El Mundo (<foreign>Radicales de Gibraltar invitan a hundir barcos españoles</foreign>) <Gibraltar Chronicle 2014-04-25>
 - (2) it appears that the plan was then that the two PSOE <indig>alcaldes</indig> from La Linea and San Roque who were due to make an appearance for a debate scheduled for today <Gibraltar Chronicle 2014-04-25>
- Identifying quotes introduced by reporting verbs, even when quotation marks are not used:
 - (3) <quote>Food for those who could afford it - officers and wealthy merchants </quote>, Bresciano **thought**. <Fall of a Sparrow 2010-11-18>

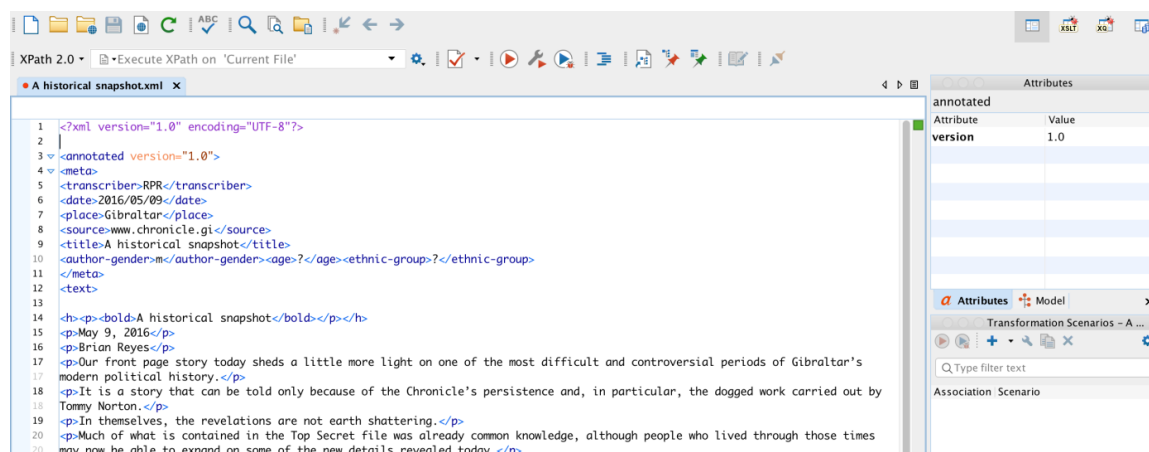


Figure 1: Oxygen xml editor: A snapshot

Regarding Novels and Short Stories, 20 further texts of 2,000 words each were needed. These were transcribed and codified applying the same criteria described above. Nonetheless, the idiosyncratic nature of each register involved a constant decision-making process which will probably only end once the final corpus is released. As an example, when transcribing one of the novels selected for inclusion in the corpus (its writer having met all the criteria to be considered a *true* Gibraltarian), we realized that much of the text corresponded to the speech of a Scottish character, speaking with a Scottish variety. The difficulty in discriminating between vernacular forms representative of Scottish or Gibraltarian English led us to discard this novel as a source.

To sum up, ICE-GBR is still in progress and the challenges ahead are considerable. However, since the section on Reportage (Press News Reports, W2C) is now fully compiled and annotated, in what follows we present a corpus-based approach to the study of language contact by comparing this section with the corresponding part of the ICE-GB component.

4. Corpus and methodology

The data under analysis here are drawn from ICE corpora (*International Corpus of English*), and are intended to represent an educated standard variety of English spoken in the regions concerned. The ICE corpora include the speech of adult (over 18 years) males and females. As noted in the previous section, in the case of ICE-Gibraltar (ICE-GBR), the speakers must also have been educated in English and have spent the last eight years in Gibraltar. ICE-GBR is currently the only source of data available on Gibraltar English, and a distinct advantage of its use here is the availability of other corpora following the same compilation guidelines, making comparisons with other varieties included in the ICE project possible. In this case the comparison is with British English (ICE-GB), and we acknowledge that there is a 30 year gap between the dates of compilation of the two corpora. From these corpora, the whole section “press news reports” (W2C) was selected for analysis, which amounts to a total of 40,000 words per variety, that is, a total of 80,000 words.

The retrieval of examples proceeded in two stages. In order to ensure a maximally exhaustive search, we first retrieved examples automatically using AntConc. The examples thus selected were analysed manually in order to select relevant examples of each of the variables under analysis (see Section 5). Since different variables were analysed, the individual searches are detailed in the analysis of each variable in Section 5.

5. Comparing press news reports in GB and GBR: Evidence of language contact?

This section explores the W2C parts of ICE-GB and ICE-GBR with the aim of discovering whether the contact with Spanish plays any role in the language of the Gibraltarian press at the morphosyntactic and lexical levels. Since the style of English and Spanish journalistic registers differ in various respects, the variables selected in this pilot study include those in which a larger difference is expected to be found. Because of space constraints, they are reduced to: (1) frequency of the passive voice, (2) type of relativizer in subject position, (3) frequency of titles and pseudo-titles, and (4) insertion of Spanish forms (code-switching).

5.1. Frequency of the passive voice

The use of the passive voice is very common in English in constructions where the agent is new information, as in (4):

- (4) *The mayor's term of office expires next month. She will be succeeded by George Hendricks.* [from Huddleston and Pullum 2002: 1444]

The passive voice as an information-rearranging strategy has been in use in English since the Middle English period (Seoane, 2006a: 361), as a result of the combination of grammatical factors, such as the levelling of inflections, and pragmatic ones, including

the tendency for topics to appear in initial position (Seoane, 2006a: 368), and it is particularly common in registers where objectivity is the main aim, as is often the case with scientific texts (see Seoane, 2006b). This is probably one of the reasons why journalists use the passive voice so frequently, an example of which would be (4) above. However, this kind of construction is not the preferred option in Spanish, which, being an inflectional language with high word order flexibility, would more typically opt for a construction such as (5):⁸

- (5) *El mandato de alcaldesa termina el próximo mes. La sucederá George Hendricks.*
her succeed.3p.sg.fut

In (5) the verb *sucedará* ‘will succeed’ is in the active voice, the subject being the following NP, *George Hendricks*. As we see, information packaging in this Spanish sentence only differs from (4) in terms of the morphosyntactic option taken (active vs. passive voice), because in both sentences the first element is a pronoun referring to the mayor, which is discourse-given.

The active voice, then, is the preferred option in spoken Spanish, and although its use in journalistic style is higher than in any other register (Rodríguez Pastor and Castro Verdala, 2013: 143), we begin with the hypothesis that it will be far more common in English press reportage than in Spanish, not least because its abuse in Spanish journalistic texts has been heavily criticized in style guides over recent decades (Rodríguez, 2001).

Thus, we have explored the 40,000 words of the W2C sections of ICE-GB and ICE-GBR by searching all possible forms of the verb BE, including both full and contracted forms. Only clear passives were included in the database; cases including an *-ed* form, which could be interpreted as an adjective, for example because it accepts intensification (see example (6)) or because it included a nearly lexicalized expression (7), were not considered actual passive constructions:

- (6) <p>Alexia said a lot of young people were **scared** of marriage and taking their vows.</p><Gibraltar Chronicle 2014-02-15>
(7) Labour 's campaign co-ordinator , Mr Jack Cunningham , said : “ The Ribble Valley by-election result shows clearly that the Government is at the end of the road and that the poll tax has to go , ” while Shadow Environment Secretary Bryan Gould said : “ The British people **are fed up** with Tory excuses and with the continuing Cabinet disarray about the future of this hated tax . ” <ICE-GB:W2C-018 #26:1>

After having excluded such unclear cases, the final tally was that shown in Table 2:

	ICE-GB	ICE-GBR
BE-PASSIVES	521	353

Table 2: Number of passive structures with *be* in the W2C sections of ICE-GB and ICE-GBR

The use of passive constructions in the press texts of ICE-GBR is much less frequent than in ICE-GB, and according to the *z-test* score for the two population proportions, the

difference is very significant.⁹ This difference, however, must not immediately be taken as a result of contact with Spanish, and we should remember that ICE-GB was released in 1988, some 26 years before ICE-GBR was even conceived. Hence, we could be witnessing a case of stylistic change. In fact, a significant decrease in the frequency of the passive in another text-type where it formerly featured extensively, namely scientific registers, has been shown to have taken place over the course of the 20th century (see, for example, Seoane, 2006b). Westin (2002: 165) finds a similar pattern in the British press, where he notes not only a reduction in the number of passives, but also in the number of subordinate clauses, as well as an increase in the use of informal language, such as interactional features (questions and imperatives) and shorter sentences. This change in the style used in some text-types has often been referred to as informalization (Leech et al., 2009: 239; Farrelly and Seoane, 2012: 395).

Informalization is defined by Farrelly and Seoane (2012: 395) as “the process whereby the distance between addresser and addressee is shortened, probably in order to make the text more engaging, accessible, and reader-friendly”. Such a process of change, which affects traditional formal registers, involves a high number of grammatical features, the exploration of which falls outside the scope of this paper, and thus will not be analysed in detail here. Nevertheless, Biber’s (1988) proposal for multidimensional analysis of registers allows for a brief exploration. Thus, if we focus on his Dimension 1 “Informational vs. Involved Production” (1988: 102), we see that one of the negative features in this dimension, which implies a higher degree of formality, is the presence of passive constructions; another is the type/token ratio, that is, more involved (meaning ‘more oral’) texts will exhibit a lower type/token ratio, while more informational (i.e. ‘less oral’) texts will exhibit a higher type/token ratio. The differences between ICE-GB and ICE-GBR regarding this feature are shown in Table 3.

ICE-GBR	# Word Types: 5224 #Word Tokens: 39357	RATIO: 7.53
ICE-GB	#Word Types: 7569 #Word Tokens: 53543	RATIO: 7.07

Table 3: Type/token ratio in the W2C sections of ICE-GB and ICE-GBR

The higher type/token ratio of ICE-GBR is also found to be significant, which does not support the hypothesis that the differences between both varieties in terms of the frequency of the passive voice are due to a higher informalization of the journalistic style in the 2010’s than in the 1980’s.

Related to informalization is colloquialization (Farrelly and Seoane, 2012), a process characterized by “a tendency for written norms to become more informal and move closer to speech” (Leech et al., 2009: 20). Within the journalistic register colloquialization manifests itself, among other ways, in an increasing use of quotations, since, as Leech et al. (2009: 20) put it, “newspaper reports now come with more direct passages of direct quotation – whether real or fictitious – than they used to”. In fact, a considerable increase in the frequency of quotations has been found in the Brown family of corpora between

the 1960's and the 1990's (see Leech et al., 2009: 128). With the aim of determining whether the difference in the use of the passive voice shown in Table 2 can be interpreted as the result of the colloquialization of the English-speaking press from the 1980's to the 2010's, we have searched the quotations used in the W2C sections of ICE-GB and ICE-GBR.

Such a search is simple to conduct, in that quotations are annotated in ICE corpora, the codes being *“*; (open quote) and *”*; (end of quote) in ICE-GB, and `<quote>` and `</quote>` in ICE-GBR. Nevertheless, and interestingly enough, we have found that in ICE-GB more quotes are opened than actually closed, as seen in (8), and on occasions some quotes are closed, but not opened (9):

- (8) `<ICE-GB:W2C-001 #63:3>`
 With the Dumfries inquiry mov<l>ing into its second month , Mr Kreindler **said** that there was nothing his group could do .
`<ICE-GB:W2C-001 #64:3>`
“We will have to wait and see if there is any substance to any of this .
`<ICE-GB:W2C-001 #65:3>`
“Meanwhile our work at the in<l>quiry will continue .
`<ICE-GB:W2C-001 #66:3>`
 What has come out so far has been very im<l>portant to us .
`<ICE-GB:W2C-001 #67:3>`
 We have been given new information into areas of cau<l>sation , and we will be able to du<l>plicate this evidence in our claim . **”**
- (9) `<ICE-GB:W2C-001 #93:5>`
 A colleague of Francisco San<l>tos , news editor of the daily El Tiempo , seized on 19 September , **commented** : : This looks very bad , very dangerous for the hostages .
”

This codification problem seen in ICE-GB is unlikely to be found in the new generation of ICE corpora, annotated in xml using the Oxygen software (as shown in Figure 1 above). Indeed, in our search for quotations in ICE-GBR there is no mismatch between the number of opening and closing quotes. That said, the number of quotations found in each corpora is shown in Table 4:

Corpus	Opening quote	Closing quote
ICE-GB	265	247
ICE-GBR	224	224

Table 4: Quotations in ICE-GB and ICE-GBR

In order to avoid any possible bias, and considering that the number of opening quotes is larger than that of closing quotes, we have taken the smaller figures of ICE-GB for the comparison between the two corpora, because in this way we make sure that we are not counting the same quotation twice, as would be the case in (8). Thus, we observe that quotations are less frequent in ICE-GBR than in ICE-GB (although the difference is not

found to be significant at $p < 0.05$, according to the *z-test* score), and thus we cannot conclude that a process of colloquialization is at work.

In summary, the frequency of passives is much lower in ICE-GBR than in ICE-GB and, having excluded the potential bias of the diachronic gap between both corpora by showing that this does not correlate with informalization or colloquialization features, the hypothesis that the Gibraltar press might be influenced by the Spanish style is still valid. Let us turn now to other features.

5.2. Relative markers

Research on relative clauses and relativization processes has generated a substantial body of work over the years, especially with regard to the distribution of relative markers within the relative clause and the factors that condition this, mainly in British English (BrE), American English (AmE) (Tottie, 1997; Guy and Bayley, 1995; Hinrichs et al., 2015) and some regional varieties (Tottie and Rey, 1997; Tottie and Harvie, 2000; Tagliamonte, 2002; Herrmann, 2003). More recently, research on relative clauses has also extended to New Englishes (Gut and Coronel, 2012; Huber, 2012; Suárez-Gómez, 2014, 2015, 2017).

Adnominal relative clauses in standard PDE are introduced by different types of relative words or relativizers, the most frequent of which are the invariable relativizer *that* (10), the pronominal relativizers represented by *wh*-words, *who* (11), *whom* (12), *whose* (13), and *which* (14), and the zero relativizer (15), traditionally distributed according to the animacy of the antecedent (human (11) vs. non-human (10)), the syntactic function of the relativizer (subject (10), object (16), complement of a preposition (12) or possessive (13)) and the type of relative clause (restrictive, RRC (10) vs. non-restrictive, NRRC (13)) (Biber et al., 1999: 608-631; Huddleston and Pullum, 2002: 1037-1057).

- (10) Many have lost one parent, some both, in the 15 years of civil war **that** have brought Angola, potentially one of the richest countries in Africa, to its knees. <ICE-GB:W2C-002 #40:2>
- (11) Other teachers believe the proposals will harm the prospects of those **who** want to specialise in traditional science subjects. <ICE-GB:W2C-002 #25:1>
- (12) Last season, he was on loan to Swansea, for **whom** he played in the European Cup. <ICE-GB:W2C-014 #28:1>
- (13) The 43-year-old man, **whose** name was withheld, will not face trial. <ICE-GB:W2C-019 #85:6>
- (14) The combination created a climate in **which** safety was not put first. <ICE-GB:W2C-007 #118:3>
- (15) Is that an irritation when you have a vague feeling **Ø** you've lent a book to somebody and you can't quite figure it out <ICE-GB:S1A-013 #92:1:A>
- (16) Publishing shares and Reuters featured on consideration of the savings **that** shareholders in the BSB satellite merger would expect to achieve. <ICE-GB:W2C-005 #61:3>

Research on the distribution of relativizers in emergent varieties of English or World Englishes, especially in Asian Englishes, shows that these varieties converge with Standard varieties of English in that *who* tends to be used as a subject with human antecedents (Levey, 2006), whereas *that* supersedes *which* in the same syntactic function with non-human antecedents, as a consequence of the general decrease of *which* at the expense of *that* (Leech et al., 2009: 227, 229). In varieties such as Indian English, however, *wh-* relativizers are clearly favoured, as is the case of *which* with non-human antecedents, irrespective of syntactic function, and this has been attributed to substratal influence, i.e. the non-reduction relativization strategy of Hindi, a language that uses pronominal relativizers to introduce the relative clause (Suárez-Gómez, 2014).

For this preliminary study on the distribution of relativizers in Gibraltar English compared to the same markers in ICE-GB, only adnominal subject relative markers in restrictive relative clauses were taken into account, that is, relative clauses introduced by a relative marker functioning as subject of the relative clause. For the search, cases of *that*, *who* and *which* were first automatically retrieved using AntConc. These searches rendered 1,545 examples which were subsequently analysed manually in order to select the relevant examples. In this process, cases of *that* as a complementizer (17) or demonstrative (18) or in combinations such as *so that* (19), and cases of relativizers which neither functioned as subjects nor introduced restrictive relative clauses (20), were excluded. This yielded a sample of 355 examples of restrictive relative clauses introduced by *that*, *who* or *which*, as Table 5 below illustrates.

- (17) The Gibraltar Government acknowledged the concerns expressed by ACI Europe and echoed the British Government's view **that** Gibraltar must be included in any EU aviation measures. <*Gibraltar Chronicle* 2016-10-10>
- (18) It becomes crystal clear that Gibraltar is not part of **that** negotiation. <*Gibraltar Chronicle* 2016-10-16>
- (19) This memorial helps us to remember all those who were killed at work but at the same time it helps us to reflect on working procedures **so that** similar tragedies are not repeated <*Gibraltar Panorama* 2014-04-29>
- (20) That is a position at odds with the British Government, **which** insists Gibraltar will be fully involved in preparations for Brexit <*Gibraltar Chronicle* 2016-11-15>

In Standard BrE, this type of relative clauses can be introduced by invariable *that* or by a *wh-* word, either *who*, if it refers back to a human antecedent, or *which*, if the antecedent is non-human. In Spanish, the invariable relativizer *que* “es el relativo de uso más general en español: puede aparecer tanto en las cláusulas especificativas como en explicativas y es la única que está capacitada para desempeñar en ambas cualquier función sintáctica (precedida, eventualmente, de la correspondiente preposición y del artículo determinado). Se trata, así pues, del nexa relativo por defecto en español”.¹⁰ (Brucart, 1999: 490); unlike English, the pronominal relativizer *quien*, only possible with human antecedents, is only possible in subject function in NRRCs (Brucart, 1999: 502).

Therefore, if language contact was operating in Gibraltar English in the choice of the relative marker functioning as subject in restrictive relative clauses, *that* would be

expected to be used more frequently than in British English, which shows variation between pronominal *wh*- words and *that*.

	Animate		Inanimate	
	<i>Who</i>	<i>That</i>	<i>Which</i>	<i>That</i>
ICE-GBR	67	2	59 (41%)	85 (59%)
ICE-GB	56	5 ¹¹	39 (46.4%)	45 (53.6%)

Table 5: Subject relativizers in RRCs in press news reports in ICE-GBR and ICE-GB

The results from Gibraltar English show a different distribution between cases of animate and inanimate antecedents, although this difference is not significant at the 0.05 level. While *who* is the default marker with animate antecedents, irrespective of the variety, *that* is more frequently used than *which* with inanimate antecedents, especially in Gibraltar English. Although we can refer to the Spanish influence to account for this slightly higher use of *that*, it is difficult to give definite conclusions because, as seen above, a peculiarity of World Englishes is that *that* supersedes *which* as subject with inanimate antecedents, also associated with a general decrease of *which* at the expense of *that* (Leech et al., 2009: 227, 229). Apart from contact, the date of compilation of the corpora analyzed is different (see section 3.1), which may also influence the results.

Summarizing, language contact with Spanish would favour the frequent use of the invariable relativizer *that*, but this is also in line with the varied literature on the distribution of relative words which agree that the *wh*-words are contracting significantly in frequency and indeed may disappear in the near future, as Schneider (1992: 446-448) forecast.

5.3. Use of titles and pseudo-titles

The use of titles in the Anglophone press is much more common than in the Spanish-speaking press, as seen in the following two examples:

- (21) Reports at the weekend claimed the Prime Minister, **Mr John Major**, would raise the issue with the United States president, **Mr George Bush**, at their summit meeting in Bermuda on Saturday, in the hope of softening US opposition to forced repatriation. <ICE-GB:W2C-019 #76:5>
- (22) **John Major** se enfrenta a los peores días de su vida. [...] Su buena relación con el presidente de Estados Unidos, **George Bush** (ambos eran los herederos grises de los carismáticos Reagan y Thatcher), su actitud sin fisuras ante el intento de golpe de estado en la antigua Unión Soviética, y su viaje a Pekín para pedir respeto a los derechos humanos, forjaron de él una interesante imagen de estadista. <El País 1992-04-10>¹²

Sentences (21) and (22) refer to two prominent politicians of the 1990's whose names are preceded by the title *Mr* in the English excerpt (taken from ICE-GB), while in the Spanish one the same public figures are referred to using their given names followed by their last names.

In addition, other nouns are also commonly used in English without a determiner and preceding a proper name, as in (23) with *President Macron*. Other cases include *lawyer*, *golfer* and *teenager* (Hundt and Kabatek, 2015) and are usually referred to as pseudo-titles (or bare noun phrases, see Kabatek and Wall, 2013):

- (23) **President Macron** has pledged to clean up French politics and public life after a series of scandals that have damaged voter confidence in their elected representatives <*The Guardian* 2017-06-20>¹³

This use cannot be said to be unattested in Romance languages, since it is recorded in Brazilian Portuguese and American-Spanish varieties, as shown in (24), and has been analysed in detail in Sáez Rivera (2013):

- (24) Eminente **juez** supremo Scarman (Argentina, <<http://veanegociador.com.ar/Inseg.htm>>, as quoted by Hundt and Kabatek 2015)

Despite being an attested use in the press of two Romance languages, the frequency of this construction is much lower than in English (Hundt and Kabatek, 2015) and, therefore, the hypothesis here is that if the Gibraltar press is somewhat influenced by its contact with Spanish, titles and pseudo-titles will be less frequent in ICE-GBR than in ICE-GB.

In order to test this hypothesis, we explored the W2C sections of both corpora, first by identifying all proper names used in the text, and then by dividing them into names preceded by a title or pseudo-title and bare proper names. The results are summarized in Table 6:

Corpus	Preceded by titles	Preceded by pseudo-titles	Proper names without any title	% of (pseudo-) titles
ICE-GBR	179 (86.5%)	28 (13.5%)	441	(207) 31.9%
ICE-GB	223 (74.6%)	76 (25.4%)	843	(299) 26.2%

Table 6: Frequency of titles and pseudo-titles in ICE-GB and ICE-GBR

As we can see, the sum of titles and pseudo-titles represents a higher percentage of NPs involving proper names in ICE-GBR than in ICE-GB (right-most column), since 31.9% of all proper names are preceded by a noun without an article in the Gibraltar data, whereas the proportion in the British press is just 26.2% (this difference is significant according to the *z-test* score). This result contrasts with those in the previous sections, because ICE-GBR appears to be more British than ICE-GB itself. Such a finding is clearly not the result of language contact or of the potential informalization of the English-speaking press (Westin, 2002, as mentioned above). Two possibilities remain. The first concerns a potential diachronic change related to a specific type of title, because pseudo-titles have been observed to experience an increase in frequency in journalistic texts (Jucker, 1992; Meyer, 2002). However, a closer look at Table 6 reveals that pseudo-titles

only represent 13.5% of all titles in ICE-GBR (example (25)), a number significantly lower than their percentage in ICE-GB (example (26)).

- (25) <p>The agreement was swiftly dismissed by **Ukip leader Nigel Farage** as a <X>”truly pathetic deal”</X>.</p><p><Gibraltar Chronicle 2016-02-20>
- (26) **Tory leader Cllr Bob Black**<l>mansaid :“ Brent Council must demonstrate to the charge-payer that it provides valuable ser<l>vices to the community. <ICE-GB:W2C-009 #54:3>

Therefore, the increasing use of pseudo-titles in global English cannot be responsible for the differences between ICE-GBR and ICE-GB. The Gibraltar press in the 2010’s shows a significant higher use of titles marking social differences, such as *Mr, Dr, Sir*, etc. than the English press in the 1980’s, as illustrated by (27) and (28):

- (27) <p>Local author and poet Jackie Anderson has organized a meeting for today for all those interested in writing. The first meeting of the Writer’s Group or Writer’s Circle will be held at the John Mackintosh Hall at 7.30pm today.</p>
- <p>The idea came to **Mrs Anderson** after the interest shown in her workshops on World Book Day.</p><p><Gibraltar Chronicle 2016-04-19>
- (28) THE Prime Minister was last night facing two crucial decisions as the shock result of the Ribble Valley by-election brought to an end his honeymoon period as **Mrs Thatcher** 's successor after only 100 days in office. <ICE-GB:W2C-018 #4:1>

Summing up, the use of titles and pseudo-titles in ICE-GB and ICE-GBR does not exhibit any evidence of contact with Spanish, nor does it show any global tendency reflecting recent historical changes regarding the use of bare noun phrases. The only explanation we can propose, then, is that of a conscious use of titles such as those in (27) and (28) as a result of the prestige of British English. It is quite possible that Gibraltarian journalists use titles with this high frequency in an attempt to sound more decidedly British, a kind of stylistic hypercorrection, while concurrently they highlight the differences between themselves and the Spanish discursive tradition of not using titles so frequently. This hypothesis, we are aware, is not uncontroversial, and much more research would be needed to confirm it.

5.4. Code-switching

Having analysed three morphosyntactic features with diverging (inconclusive) results regarding the influence of Spanish as a contact language with Gibraltarian English (passives, relative markers and use of (pseudo-)titles), this section explores the lexis used in the news reports included in ICE-GBR, with the aim of finding a pattern in the use of Spanish words and expressions. Code-switching is one of the common results of language contact (see section 2 and also Thomason and Kaufmann, 1988; Thomason, 2001; Bolton, 2006) and is one of the idiosyncratic features of Gibraltarian speech, as reported in Kramer (1986), Moyer (1993) and subsequent studies, because, as Gal (1988: 247) says:

code-switching practices are symbolic creations concerned with the construction of the self and the other within a broader political, economic and historical context which reflect the way people respond symbolically to relations of domination between groups within the state, and how they understand this historic position and identity within a world capitalist system structure around dependency and unequal development (Gal, 1988: 247).

These “symbolic creations” are expected to be much more common in spoken registers than in written ones, but of all the types of switches identified by Van Hout and Muysken (1995) we have found that *insertions* do appear in our corpus, although at a low frequency. Of the total of 40,000 words in the W2C section of ICE-GBR, six Spanish forms are inserted into the English texts, as shown in Table 7:

Spanish form	Number of occurrences
<i>Sr</i>	5
<i>Snr</i>	3
<i>alcalde(s)</i>	2
<i>mano a mano</i>	1
<i>a mí me pertenece</i>	1
<i>Pueblo</i>	1

Table 7: Code-switching cases in ICE-GBR

Interestingly enough, the two most common Spanish forms are titles, namely abbreviations of Spanish *señor* ‘mister’, which exclusively precede names of Spanish-speaking people (*Sr Rajoy*, *Sr Margallo*, both Spanish politicians, and *Sr Morejon-Pazmiño*, an Ecuadorian diplomat).¹⁴ None of these Spanish names is ever preceded by its English counterpart *Mr*, whose use is restricted to non-Spanish people, such as *Mr Cameron*, *Mr Clinton*, *Mr Cooper* and also *Mr Guerrero*, who, despite having a Spanish surname, is a Gibraltarian citizen, the CEO of the Gibraltar Tourist Board. The shortening *Snr* is also worth mentioning, since it is not a Standard Spanish abbreviation (as noted by the *Real Academia de la Lengua Española*).¹⁵ Its use stands alongside that of *Sr*, even in the same text (e.g. *Snr Diego Morejon* and *Sr Morejon*), but it seems to be restricted to one author, Fabian Picardo, Gibraltar’s chief minister.

The word *alcalde(s)* is used to refer to Spanish mayors, although the use of the English word is also possible when referring to Spanish politicians, as seen in (29) and (30):

- (29) <p>Whatever evolution that moot invitation took, if at all, it appears that the plan was then that the two PSOE <indig>**alcaldes**</indig> from La Linea and San Roque who were due to make an appearance for a debate scheduled for today, Saturday, asked for this to be changed for diary reasons.</p><p><Gibraltar Chronicle 2014-04-25>

- (30) <p>Speaking later to Spanish journalists Mr Feetham said he was saddened that he was invited to sit in a debate with the mayors and the La Linea mayor had then pulled out <quote>"not to offend Mr Picardo...who was in the Caymans anyway"</quote>.</p><p><i>Gibraltar Chronicle 2014-04-25</i></p>

The reason for this code-switching in (29) is, therefore, not justified by a linguistic gap, but by its expressive function (Poplack, 1980). Expressivity also seems to be the reason behind the use of the sequences *mano a mano* and *a mí me pertenece*, as shown in (31) and (32):

- (31) <p>Locally nothing changed on the running <indig>**mano a mano**</indig> between Fabian and Daniel, the latter who rose to Leader of the Opposition whilst his predecessor was in June to arise as Sir Peter.</p><p><i>Gibraltar Chronicle 2014-01-23</i></p>
- (32) <p>Mr Feetham describes his project as <quote>"seeking to break the mould of Gibraltar politics".</quote></p><p><quote>"I want to break from promising people things, of promising everything to everyone. Gibraltar cannot continue with the culture of <punctuation>'</punctuation><indig>**a mi me pertenece**</indig><punctuation>'</punctuation>. It will catch up on us,"</quote> he says warning that if Gibraltar were hit by difficulties the issues he is warning about will become critical.</p><p><i>Gibraltar Chronicle 2014-01-23</i></p>

In sentences (31) and (32), two Spanish expressions are inserted in the discourse, although their equivalent English forms would suit the communicative purposes perfectly: *hand in hand* and *it belongs to me*. The reason for code-switching, then, has to do with a high degree of expressiveness and the construction of a Gibraltarian identity, in the sense described by Gal (1988), above.

Summing up, these cases of code-switching constitute strong evidence for the influence of Spanish on the Gibraltarian press. Although the native code-switching variety *Yanito* exists mainly in the spoken mode, Gibraltarian speakers seem to have a tendency to insert occasionally Spanish words and expressions into written registers, such as press news reports.

6. Summary and conclusions

This paper has presented a preliminary study of the effects of language contact in Gibraltar by exploring the reportage sections of two components of the *International Corpus of English*, namely ICE-GB (Great Britain) and ICE-GBR (Gibraltar). The ICE project was born in the 1980's with the idea of providing comparable one-million-word corpora of different varieties of English, and the first member of the family was ICE-GB, released in 1988. The team to which the authors belong was commissioned in 2014 to compile the ICE-GBR, and currently 100,000 words have been completed. For this preliminary study only 40,000 words were used from each corpus, namely those including press news reports. With the aim of investigating whether the Gibraltarian press variety

is influenced to some extent by its contact language, Spanish, four different features were scrutinized, the frequency of the passive voice, the type of relativizer in subject position, the use of (pseudo-)titles, and code-switching. The results are summarized in Table 8.

Feature	Evidence of contact with Spanish	Other possible explanation
Frequency of the passive voice	Possibly	No. Informalization and colloquialization have been ruled out.
Type of relativizer in subject position	Difficult to determine	Global tendency in World Englishes.
Use of (pseudo-)titles	No	Prestige of British English?
Code-switching	Yes	No

Table 8: Summary of results

It was hypothesized that, if Spanish does influence the Gibraltarian press, then the frequency of the passive voice would be lower in ICE-GBR than in ICE-GB. This was indeed seen to be the case, and was statistically significant. Further tests were conducted to rule out a potential diachronic change (owing to the 30-year gap between the compilation of the two corpora), but the exploration of the type/token ratio (in line with Biber, 1988) and the frequency of quotations (following Leech et al., 2009) suggest that ICE-GBR does not exhibit features related to informalization or colloquialization.

The lower frequency of the passive voice in the Gibraltarian press is, therefore, not the result of either of these two processes, which implies that contact with Spanish may possibly play a role in this morphosyntactic feature.

Regarding the type of relativizer, the hypothesis was that the overwhelming presence of Spanish *que* would favour a similar use of English *that*. This was found to be the case with inanimate antecedents (but not with animate ones), although asserting that this is the result of the Spanish influence is premature, since other varieties of English worldwide have witnessed the same development recently, especially in the low frequency of *which* with inanimate antecedents, and this seems to be the trend in native varieties of English, as reported by Leech et al. (2009: 227, 229).

The third variable studied was the use of titles (e.g. *Mr, Dr, Sir*) and pseudo-titles (e.g. *Tory leader, therapist, etc.*), with the hypothesis that the Spanish influence would render a much lower frequency of these items in ICE-GBR than in ICE-GB. The results, in this case, clearly go against the hypothesis, since the Gibraltarian press includes (pseudo-)titles with a significantly higher frequency than ICE-GB. This constitutes another piece of evidence for the lack of informalization and colloquialization of this text-type in ICE-GBR; it was also suggested that the reason for this unexpected result could be rooted in a voluntary attempt to distance the text from Spanish discourse traditions and, concomitantly, a desire to invoke the higher prestige of British English.

Finally, cases of code-switching were also examined, since the combination of English and Spanish is common in the spoken mode (and the source for the Gibraltar variety *Yanito*, see Section 2). The press news reports included in ICE-GBR exhibit several cases of insertion of Spanish forms for expressive purposes, such as the title *Sr*, the noun *alcalde*, and constructions such as *mano a mano*. This final variable is, undoubtedly, the most conspicuous piece of evidence for the contact between English and Spanish in Gibraltar.

All in all, we have shown that the compilation of comparable corpora constitutes a necessary tool for the study of variation in general, and for the analysis of language contact phenomena in particular. Although the compilation of ICE-GBR is ongoing, this preliminary study demonstrates that, despite the fact that the English spoken in Gibraltar exhibits a strong exonormative allegiance to British English, the language used by the Gibraltar press exhibits evidence of the influence of Spanish, as well as some idiosyncratic features which require further analysis. The findings in this study will have to be compared with the exploration of more text-types and, finally, with the whole ICE-GBR corpus, which will provide a comprehensive picture of Gibraltar English at the beginning of the 21st century.

Notes

* The authors gratefully acknowledge the Spanish Ministry of Economy and Competitiveness (grants FFI2014-53930-P and FFI2014-51873-REDT). Thanks are also due to two anonymous reviewers for their helpful comments.

1. MG Sanchez, a Gibraltar writer, reflects this idea in *The Escape Artist*: “Most Gibraltarians have [...] mixed bloodlines. This is because most of us are the sons of immigrants who came to Gibraltar from other parts of the Mediterranean from the mid-1750s onwards. In fact, in the nineteenth-century everybody used to call us ‘mongrels’ because we were neither Spanish nor British, but a strange composite of the two with a bit of Genoese, Maltese, Irish, Portuguese and Jewish through into the melting pot as well” (2013: 13).

2. See <<http://view0.webs.uvigo.es/ice-gibraltar>> for further details, last access 25 September 2017.

3. Official census:<<https://www.gibraltar.gov.gi/new/gibraltar-census-history>>, last access 25 September 2017.

4. The issue of identity underlies most references to the linguistic situation of Gibraltar. Kellerman (2001: 411) defines Gibraltar as a region in search of an identity, a language and a culture. For Moyer (1993: 236), “Gibraltarians avoid identifying themselves with a given group, and at the same time they are affirming a local identity”. This is also clearly shown in the novels of MG Sanchez, one of the most prolific local writers. In *The Escape Artist* we find the following excerpts: “‘I am Gibraltarian,’ he said in a tone which made it clear that he was no longer kidding”. (Sanchez, 2013: 5) and “No, I’m not Spanish. I’m from Gibraltar, mate” (Sanchez, 2013: 14). Finally, Jennifer Ballantine in a recent interview states: “I was away from Gibraltar for a period of years, and from when I left in 1994 to when I came back in 2007, in that period of time this question of identity had somehow erupted in a way that I hadn’t heard before, certainly not when I left initially. I suspect that the tercentenary commemoration of the taking of Gibraltar in 2004 would have focused everybody’s minds towards the fact that it was three hundred years

since Gibraltar was taken from Spain, and this must mean something; at that time a number of publications were produced which aimed to look at the question of Gibraltar” (Seoane, 2017: 222).

5. This age-grading pattern (Weston, 2011; Seoane et al., 2016) is also reinforced by MG Sanchez, who agrees with the assumption that the choice of language is directly related to age (personal communication). Nevertheless, according to Jennifer Ballantine (cf Seoane, 2017: 223) we cannot generalize. In her opinion, “I think that social class and educational background has a certain amount of influence in this question,” although she agrees that, before WWII, Spanish was very widespread as shown by the existence of Gibraltarian newspapers other than the *Gibraltar Chronicle*, such as *El Calpense* or *El Anunciador*, both written in Spanish produced for a local readership. Nowadays, however, the youngest speakers are mostly monolingual in English (“I think that English is becoming the main language for this generation, and code switching is probably very difficult; only certain terms enter their register but not every term that perhaps an older generation would be able to use”). The closure of the Instituto Cervantes in Gibraltar in 2015 is further evidence for this increasing gap between the Spanish language and the Gibraltarian population.

6. As acknowledged by MG Sanchez, “[t]o understand modern-day Gibraltar you have to understand this sense of hybridity and all the apparent craziness that comes with it – blond-haired people with Spanish surnames, red pillar boxes beside *chiringuitos* serving *gambas a la plancha*, British bobbies speaking Spanish, et cetera, et cetera” (Seoane, 2016: 255).

7. Even as this article is being written, changes are taking place regarding ICE’s website and its coordination. The former coordinator, Gerald Nelson (Chinese University of Hong Kong), has just been succeeded by Marianne Hundt (Univ. of Zürich) and a board is in the process of being constituted, as announced at a recent ICAME conference (Kirk and Nelson 2017). Likewise, ICE’s webpage is in the process of being moved from its present hosting site to an institutional URL at the University of Zürich.

8. Spanish has a passive construction with the pronoun *se* (e.g. *Se venden limones* ‘Lemons are sold’; see Mendikoetxea (1999: 1631-1722)), which falls out of our study here.

9. The *z-test* score was obtained using the online calculator available at <<http://www.socscistatistics.com/tests/ztest/Default2.aspx>>, last access 25 September 2017.

10. The invariable relativizer *que* “is the most common relativizer in Spanish: it can appear both in restrictive and non-restrictive relative clauses, and it is the only relativizer that can play any syntactic function (preceded, if needed, by the corresponding preposition and the definite article). ... It is, therefore, the default relativizer in Spanish” [our translation].

11. Here examples with antecedents such as *army* or *government* were found, as in *When Beirut and its environs will be declared officially free of militias and under the control of a reunited Lebanese army that has, for the first time, the germ of sectarian balance* <ICE-GB:W2C-010 #77:2>. They refer to a group of people, therefore they have an animate antecedent, but they also allow relative clauses introduced by *which*.

12. This sentence was taken from:

<http://elpais.com/diario/1992/04/10/internacional/702856801_850215.html>, last access 25 September 2017.

13. From <<https://www.theguardian.com/world/2017/jun/20/french-defence-minister-sylvie-goulard-resigns-after-a-month>>, last access 25 September 2017.

14. Spelled *Morejon-Pazmino* in the Gibraltarian press.

15. See <<http://buscon.rae.es/dpd/apendices/apendice2.html>> for a full list of accepted Spanish abbreviations, last access 25 September 2017.

References

- Ballantine, Sergius J. (1983): "A study of the effects of English-medium education on initially monoglot Spanish-speaking Gibraltar children". MA dissertation, University of Valencia.
- Ballantine, Sergius J. (2000): "English and Spanish in Gibraltar: Development and characteristics of two languages in Gibraltar". *Gibraltar Heritage Journal*, 7: 115–124.
- Biber, Douglas (1988): *Variation Across Speech and Writing*. Cambridge: Cambridge University Press.
- Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad and Edward Finegan (1999): *Longman Grammar of Spoken and Written English*. Harlow: Longman.
- BNC = *The British National Corpus*, version 3 (BNC XML Edition) (2007): Distributed by Oxford University Computing Services on behalf of the BNC Consortium. URL: <<http://www.natcorp.ox.ac.uk/>> [accessed 10 June 2017].
- Bolton, Kingsley (2006): "World Englishes Today". In Braj B. Kachru, Yamuna Kachru and Cecil L. Nelson, eds. *The Handbook of World Englishes*. Oxford: Blackwell, 240-269.
- Brown = *A Standard Corpus of Present-Day Edited American English, for use with Digital Computers*. Compiled by W. Nelson Francis and Henry Kučera (Brown University, Providence, Rhode Island).
- Brucart, José Ma. (1999): "La estructura del sintagma nominal: las oraciones de relativo". In Ignacio Bosque and Violeta Demonte, coords., *Gramática descriptiva de la lengua española. Vol. I Sintaxis básica de las clases de palabras*. Madrid: Espasa Calpe, 395-522.
- Cal Varela, Mario (1996): "Hacia una concepción prototípica de comunidad de habla: Gibraltar". *Atlantis*, XVIII(1–2): 37–52.
- Cal Varela, Mario (2001): *Algunos aspectos sociolingüísticos del inglés gibraltareño: Análisis cuantitativo de tres variables a nivel fónico*. Santiago de Compostela: Servizo de Publicacións, Universidade de Santiago de Compostela.
- Cavilla, Manuel ([1978] 1990): *Diccionario Yanito*. 2nd edition. Gibraltar: MedSUN.
- Census of Gibraltar (2012): <https://www.gibraltar.gov.gi/new/census> [accessed 15 June 2017].
- Davies, Mark (2008-): *The Corpus of Contemporary American English: 520 Million Words, 1990-present*. Available at <<http://corpus.byu.edu/coca/>> [accessed 15 May 2017].
- Davies, Mark (2013): *Corpus of Global Web-Based English: 1.9 Billion Words from Speakers in 20 Countries*. Available at <<http://corpus.byu.edu/glowbe/>> [accessed 15 May 2017].
- Farrelly, Michael and Elena Seoane (2012): "Democratisation". In Terttu Nevalainen and Elizabeth C. Traugott, eds., *The Oxford Handbook of the History of English*. Oxford: Oxford University Press, 392-401.
- Gal, Susan (1988): "The political economy of code choice". In Monica Heller, ed., *Code-Switching: Anthropological and Sociolinguistic Perspectives*. Berlin: De Gruyter, 245-263.
- García-Castro, Laura (2017): "'I remember to read a statement from somebody on net': clausal complementation in L2 varieties of English". Paper presented at the *Santiago-Leuven-Edinburgh seminar on Grammatical Variation and Change in English*. University of Leuven, 4-5 April 2017.
- Greenbaum, Sidney (1988): "A proposal for an international computerized corpus of English". *World Englishes*, 7(3): 315.
- Gut, Ulrike and Lilian Coronel (2012): "Relatives worldwide". In Marianne Hundt and Ulrike Gut, eds., *Mapping Unity and Diversity Worldwide: Corpus-based Studies of New Englishes*. Amsterdam: John Benjamins, 215-242.

- Guy, Gregory R. and Robert Bayley (1995). "On the choice of relative pronouns in English". *American Speech*, 70(2): 148-172.
- Herrmann, Tanja (2003): *Relative Clauses in Dialects of English: A Typological Approach*. Freiburg, Germany: University of Freiburg dissertation.
- Hinrichs, Lars, Benedikt Szmrecsanyi and Axel Bohmann (2015): "Which-hunting and the Standard English relative clause". *Language*, 91(4): 806-836.
- Huber, Magnus (2012): "Syntactic and variational complexity in British and Ghanaian English relative clause formation in the written parts of the International Corpus of English". In Bernd Kortmann and Benedikt Szmrecsanyi, eds., *Linguistic Complexity*. Berlin and Boston: De Gruyter, 218-242.
- Huddleston, Rodney and Geoffrey K. Pullum (2002): *The Cambridge Grammar of the English Language*. Cambridge: Cambridge University Press.
- Hundt, Marianne (2015): "World Englishes". In Douglas Biber and Randi Reppen, eds., *The Cambridge Handbook of English Corpus Linguistics*. Cambridge: Cambridge University Press, 381-400.
- Hundt, Marianne and Johannes Kabatek (2015): "Pseudo-titles, global language contact and discourse traditions". Paper presented at the 48th Meeting of the Societas Linguistica Europaea, Leiden, Netherlands, 2-5 September 2015.
- Jucker, Andreas (1992): *Social Stylistics. Syntactic Variation in British Newspapers*. Berlin: De Gruyter.
- Kabatek, Johannes and Albert Wall, eds. (2013): *New Perspectives on Bare Noun Phrases in Romance and Beyond*. Amsterdam: John Benjamins.
- Kellerman, Anja (2001): *A New New English. Language, Politics and Identity in Gibraltar*. Heidelberg: Heidelberg Schriften zur Sprache und Kultur.
- Kirk, John and Gerald Nelson (2017): "The International Corpus of English (ICE) Project: The next 25 years". Plenary Lecture delivered at the 38 ICAME Conference. Prague, Czech Republic, 24-28 May 2017.
- Kramer, Johannes (1986): *English and Spanish in Gibraltar*. Hamburg: Helmut Buske Verlag.
- Le Page, Robert B. and André Tabouret-Keller (1985): *Acts of Identity: Creole-based Approaches to Language and Ethnicity*. Cambridge: Cambridge University Press.
- Leech, Geoffrey, Marianne Hundt, Christian Mair and Nicholas Smith (2009): *Change in Contemporary English: A Grammatical Study*. Cambridge: Cambridge University Press.
- Levey, David (2008): *Language Change and Variation in Gibraltar*. Amsterdam and Philadelphia: John Benjamins.
- Levey, David (2015): "Gibraltar English". In Daniel Schreier, Peter Trudgill, Edgar W. Schneider and Jeffrey P. Williams, eds., *The Lesser-Known Varieties of English*. Vol 2. Cambridge: Cambridge University Press, 51-69.
- Levey, Stephen (2006): "Visiting London relatives". *English World Wide*, 27: 45-70.
- Lipski, John (1986): "Sobre el bilingüismo anglo-hispánico en Gibraltar". *Neophilologische Mitteilungen*, LXXXVIII: 414-427.
- LOB Corpus = Lancaster-Oslo/Bergen Corpus. Compiled by Geoffrey Leech (Lancaster University), Stig Johansson (University of Oslo) and Knut Hofland (University of Bergen).
- Loureiro-Porto, Lucía (2016): "(Semi-)modals of necessity in Hong Kong and Indian Englishes". In Elena Seoane and Cristina Suárez-Gómez, eds., *World Englishes: New Theoretical and Methodological Considerations (Varieties of English Around the World, G57)*. Amsterdam: John Benjamins, 143-172.

- Loureiro-Porto, Lucía (2017): "ICE vs. GloWbE: Big data and corpus compilation". *World Englishes*, 36(3): 448-470.
- Mendikoetxea, Amaya (1999): "Construcciones con "se": medias, pasivas e impersonales". In Ignacio Bosque and Violeta Demonte, coords. *Gramática descriptiva de la lengua española. Vol. II Las construcciones sintácticas fundamentales. Relaciones temporales, aspectuales y modales*. Madrid: Espasa Calpe, 1631-1722.
- Meyer, Charles F. (2002): "Pseudo-titles in the press genre of various components of the International Corpus of English". In Randi Reppen, Susan M. Fitzmaurice and Douglas Biber, eds., *Using Corpora to Explore Linguistic Variation*. Amsterdam: John Benjamins, 147-166.
- Moyer, Melissa G. (1993): *Analysis of Code-Switching in Gibraltar*. Barcelona: Publicacions de la Universitat Autònoma de Barcelona.
- Moyer, Melissa G. (1998): "Bilingual conversation strategies in Gibraltar". In Peter Auer, ed., *Code-Switching in Conversation: Language, Interaction and Identity*. New York: Routledge, 215-234.
- Mukherjee, Joybrato and Marco Schilk (2012): "Exploring variation and change in New Englishes. Looking into the *International Corpus of English (ICE)* and beyond". In Terttu Nevalainen and Elizabeth Closs Traugott, eds., *The Oxford Handbook of the History of English*. Oxford: Oxford University Press, 189-199.
- Nelson, Gerald (2006): "World Englishes and corpora studies". In Braj Kachru, Yamuna Kachru and Cecil L. Nelson, eds., *The Handbook of World Englishes*. Oxford: Blackwell, 733-750.
- Poplack, Shana (1980): "Sometimes I'll start a sentence in Spanish Y TERMINO EN ESPAÑOL: toward a typology of code-switching". *Linguistics*, 18: 581-618.
- Rissanen, Matti (2014): "On English historical corpora, with notes on the development of adverbial connectives". In Alejandro Alcaraz-Sintes and Salvador Valera-Hernández, eds., *Diachrony and Synchrony in English Corpus Linguistics (Linguistic Insights 181)*. Bern: Peter Lang, 109-139.
- Rodríguez, José Manuel (2001): "Errores comunes en el lenguaje periodístico: Invasión pasiva". *Revista Latinoamericana de Comunicación CHASQUI* 76. Available at <<http://www.redalyc.org/pdf/160/16007609.pdf>> [accessed 10 June 2017].
- Rodríguez Pastor, César Luis and Verónica Castro Verdala (2013): "Voz pasiva perifrástica. Una propuesta didáctica". *Foro de Profesores de E/LE*, 9: 143-151.
- Sáez Rivera, Daniel M. (2013): "Causes and conditions for the lack of the definite article in American-Spanish headlines". In Johannes Kabatek and Albert Wall, eds., *New Perspectives on Bare Noun Phrases in Romance and Beyond*. Amsterdam: John Benjamins, 157-188.
- Sanchez, Mark G. (2013): *The Escape Artist. A Gibraltarian Novel*. United Kingdom: Rock Scorpion Books.
- Schaub, Steffen (2016): "The influence of register on noun phrase complexity on varieties of English". In Christoph Schubert and Christina Sanchez-Stockhammer, eds., *Variational Text Linguistics: Revisiting Register in English*. Berlin: De Gruyter, 251-270.
- Schneider, Edgar (1992): "Who(m)? Constraints on the loss of case marking of *wh*-pronouns in the English of Shakespeare and other poets of the Early Modern English period". In Matti Rissanen, Ossi Ihalainen, Terttu Nevalainen and Irma Taavitsainen, eds., *History of Englishes: New Methods and Interpretations in Historical Linguistics*. Berlin: De Gruyter, 437-452.
- Schneider, Edgar (2007): *Postcolonial English: Varieties around the World*. Cambridge: Cambridge University Press.

- Seoane, Elena (2006a): "Information structure and word order change: The passive as an information rearranging strategy in the history of English". In Ans van Kemenade and Bettelou Los, eds., *The Handbook of the History of English*. Oxford: Blackwell, 360-391.
- Seoane, Elena (2006b): "Changing styles: On the recent evolution of scientific British and American English". In Christiane Dalton-Puffer, Dieter Kastovsky and Nikolaus Ritt, eds., *Syntax, Style and Grammatical Norms: English from 1500-2000*. Bern: Peter Lang, 191-211.
- Seoane, Elena (2016): "Telling the true Gibraltarian story: an interview with Gibraltarian Writer M.G. Sanchez". *Revista Alicantina de Estudios Ingleses*, 29: 251-258.
- Seoane, Elena (2017): "A Gibraltar in the making: Interview with Dr. Jennifer Ballantine Perera". *Canarian Journal of English Studies*, 75: 217-225.
- Seoane, Elena and Cristina Suárez-Gómez (2013): "The expression of the perfect in East and South-East Asian Englishes". *English World-Wide*, 34(1): 1-25.
- Seoane, Elena, Cristina Suárez-Gómez and Lucía Loureiro-Porto (2016): "The ICE project looks at Iberia: The International Corpus of Gibraltar English". Paper presented at the *40th AEDEAN Conference*. Universidad de Zaragoza, 8-10 November 2016.
- Suárez-Gómez, Cristina (2012): "English in contact with other European languages (Italian, Spanish, Slavic)". In Alexander Bergs and Laurel J. Brinton, eds., *Historical Linguistics in English: An International Handbook*. Vol 2. Berlin: De Gruyter, 1738-1753.
- Suárez-Gómez, Cristina (2014): "Relative clauses in Asian English". *Journal of English Linguistics*, 42(3): 245-268.
- Suárez-Gómez, Cristina (2015): "*The places where English is spoken*: adverbial relative clauses in World Englishes". *World Englishes*, 34(4): 620-635.
- Suárez-Gómez, Cristina (2017): "The role of transparency and language contact in the structural nativization of relative clauses in New Englishes". *English World-Wide*, 38(2): 212-238.
- Tagliamonte, Sali (2002): Variation and change in the British relative marker system. In P. Poussa, ed., *Dialect Contact on the North Sea Littoral*. Munich: Lincom Europa, 147-165.
- The ICE Project. 2009. The design of ICE corpora. Available online at <<http://ice-corpora.net/ice>> [accessed April 2016].
- Thomason, Sarah G. (2001): *Language Contact: An Introduction*. Edinburgh: Edinburgh University Press.
- Thomason, Sarah G. and Terrence Kaufmann (1988): *Language Contact, Creolization, and Genetic Linguistics*. Berkeley and Los Angeles: University of California Press.
- Tottie, Gunnel (1997): "Relatively speaking. Relative marker usage in the British National Corpus". In Terttu Nevalainen and Leena Kahlas-Tarkka, eds., *To Explain the Present: Studies in the Changing English Language in Honour of Matti Rissanen*. Helsinki: Societe Neophilologique, 465-481.
- Tottie, Gunnel and Michael Rey (1997): "Relativization strategies in earlier African American Vernacular English". *Language Variation and Change*, 9(2): 219-247.
- Tottie, Gunnel and Dawn Harvie (2000): "It's all relative: Relativization strategies in early African American English". In Shanna Poplack, ed., *The English History of African American English*. Oxford: Blackwell, 198-230.
- Van Hout, Roeland and Pieter Muysken (1995): "Alternation, insertion, congruent lexicalizations: Corpus based and structural approaches". *Summer School on Code-Switching and Language Contact. Ljouwert/Leeuwarden, 14-17 September 1994*. Ljouwert/Leeuwarden: Fryske Akademy, 302-306.
- Westin, Ingrid (2002): *Language Change in English Newspaper Editorials*. Amsterdam: Rodopi.

- Weston, Daniel (2011): "Gibraltar's position in the Dynamic Model of Postcolonial English". *English World-Wide*, 32(3): 338-367.
- Weston, Daniel (2013): "Code-switching variation in Gibraltar". *International Journal of Bilingualism*, 17(1): 3-22.
- Weston, Daniel (2015): "The lesser of two evils: Atypical trajectories in English dialect evolution". *Journal of Sociolinguistics*, 19(5): 671-687.

Contact-induced variation in clausal verb complementation: the case of REGRET in World Englishes¹

Raquel P. Romasanta
Universidade de Vigo
rapereira@alumnos.uvigo.es

It has been argued that in language contact situations both transfer processes from the substrate languages (Thomason, 2008) and cognitive effects derived from the language contact situation itself (Schneider, 2012, 2013) can constitute important catalysts for language variation and change. Regarding the verbal complementation system, Steger and Schneider (2012: 172), for example, notice a preference for finite patterns over non-finite structures in World Englishes (WEs), that is, a preference for more explicit forms (hyperclarity and isomorphism). On the contrary, Schneider's study (2012) does not confirm such a preference for more explicit forms in WEs in the competition between finite and non-finite patterns. This article intends to shed some light on the differences between the distribution of finite and non-finite complementation patterns in WEs by exploring the complementation profile of the verb REGRET in two metropolitan varieties, British and American English, and comparing them to three geographically distant varieties with different substrate languages, historical contexts, and degrees of language contact: on the one hand, two ESL varieties, Hong Kong English and Nigerian English, and on the other, one ESD variety, Jamaican English, where contact is more pronounced. The main aim of this paper is, therefore, to investigate whether potential differences in the verbal complementation systems between varieties of English are product of cognitive processes derived from the language contact situation, a matter of transfer-induced change, or a combination of both.

Keywords: clausal complementation, language contact, World-Englishes, substrate influence, cognitive effects, transfer



1. Introduction

The English clausal complementation system has received a great deal of attention in the literature, not least in diachronic studies, where the emphasis has typically been on supranational or native varieties. It has been seen to have undergone a series of major and minor changes over the centuries, these known as the Great Complement Shift (cf. Rohdenburg, 2006: 143) or, in Sapir's terms (1921), as a drift (cf. Fanego, 2007: 162, see also, among others, Warner, 1982; Fischer, 1988, 1989; Fanego, 1990, 1992, 1996a, 1996b, 1998, 2004a, 2004b, 2010, 2016; Rohdenburg, 1995, 2006, 2014; Rudanko, 1998, 2000, 2011; Miller, 2002; Los, 2005; Vosberg, 2006; and De Smet, 2008, 2009, 2010, 2013, 2014). However, further work on variation in clausal complementation is still needed, especially in Present-Day English "where comparatively little work has been done" (Fanego, 2007: 161), and also in World Englishes, since, as Schneider (2007: 86) argues, "a classic example [of innovations in varieties in phase 4, nativization] is the complementation patterns which verbs and also adjectives typically enter". Thus far, studies on verbal complementation in World Englishes (WEs) are scarce and have centered mainly on nominal complementation with focus on ditransitive verbs and Transfer-Caused-Motion (TCM) constructions (see Olavarria de Ersson and Shaw, 2003; Mukherjee and Hoffman, 2006; Mukherjee and Schilk, 2008, 2012; Mukherjee and Gries, 2009; Bernaisch, 2013; Deshors, 2014); as for clausal complementation, only quantitative studies on the competition between gerunds and infinitives are currently available (Deshors, 2015; Deshors and Gries, 2016).

The complementation system in these studies is seen as being innovative and indeed divergent from one variety of English to another (cf. Mukherjee and Hoffmann, 2006; Schneider, 2007). Regarding differences between varieties, it has been argued that in language contact situations both transfer processes from the substrate languages (Thomason, 2008) and cognitive effects derived from the language contact situation itself (Williams, 1987; Schneider, 2012, 2013) can constitute important catalysts for language variation and change. In the case of English as a contact language, increased isomorphism can at times be seen in the verbal complementation system: Steger and Schneider (2012: 172), for example, hypothesize that WEs should show a preference for finite patterns over non-finite structures, that is, a preference for more explicit forms. In confirming their hypothesis, they also find that "prototypically non-finite verbs display instances of not only intermediate but even finite complementation in the corpus [ICE]" (Steger and Schneider, 2012: 179). However, another study on clausal complementation in L2 varieties (Hong Kong English, East Africa English, Indian English, and Singapore English) by Schneider (2012) has shown that the hypothesis that WEs tend towards isomorphism and increased explicitness is not confirmed. Schneider (2012: 80) attributes the lack of isomorphism in clausal complementation to the fact that he focuses on high-frequency verbs, such as BELIEVE, PROMISE, and WISH, which might increase the "stability of transmission" (Schneider, 2012: 80). Therefore, in the present study I will analyze a low-frequency verb, REGRET, which allows for variation between finite and non-finite patterns and should not, in principle, show such stability of transmission.

Because the previously mentioned studies (Steger and Schneider, 2012; Schneider, 2012) do not consider substrate languages as a possible factor influencing the complementation system of non-native varieties of English, the main aim of this article is to investigate whether potential differences in the complementation systems between supranational varieties of English, English as a Second Language (ESL) varieties, and English as a Second Dialect (ESD) varieties are, (i) the product of cognitive processes derived from the language contact situation, (ii) a matter of transfer-induced change, that is, influence of the substrate language(s), or (iii) a combination of both. With this in mind, I will examine the complementation profile of the verb *REGRET* in two metropolitan varieties, American (AmE) and British (BrE) English, and compare these to three geographically distant varieties with different substrate languages, historical contexts, and degrees of language contact: on the one hand, two ESL varieties, Hong Kong English (HKE) and Nigerian English (NigE), and on the other, one ESD variety, Jamaican English (JamE), where contact is more pronounced (cf. section 2.3 below). The low frequency of use of the verb *REGRET* meant that the *International Corpus of English* (ICE) was too small as a data source and I therefore used the *Corpus of Global Web-Based English* (GloWbE, Davies, 2013). This corpus provides a vast amount of material on a good range of varieties of English, allowing both the study of low frequency structures such as clausal complementation and also comparisons between different varieties of English.

In section 2, below, I present a brief account of previous literature on the English clausal complementation system in general, the verb *REGRET* in particular, and on language contact phenomena and its repercussions in WEs. In section 3, the methodology is described, followed by the results and analysis. Finally, I summarize the main conclusions in section 5.

2. Background

2.1. The English clausal complementation system and the verb *REGRET*

The English complementation system has undergone a huge restructuring over the centuries, commonly referred to as the Great Complement Shift (Rohdenburg, 2006: 143). Two of the most notable of these changes are the spread of the infinitive at the expense of finite clauses (see Rohdenburg, 1995) and the establishment of the gerund as a second type of non-finite complement alongside infinitives after it developed verbal features during Late Middle English (Fanego, 1996a, 2004b).

For Present-Day English, reference grammars (Quirk et al., 1985; Biber et al., 1999; Huddleston and Pullum et al., 2002) consider four major structural types of complement clauses, organized according to their internal structure and complementizer: *wh*-clauses (or interrogatives in Huddleston and Pullum et al., 2002 terms), *to*-infinitive-clauses, *-ing*-clauses or gerunds, and *that*-clauses (Huddleston and Pullum et al., 2002: 951), see examples (2) to (6) below.

As for the verb *REGRET*, Framenet² recognizes six different possible complementation patterns, namely noun phrase, as in example (1), *wh*-clauses, as in example (2), *to*-

infinitive-clauses, as in example (3), *-ing*-clauses, which may occur with or without an overt subject,³ see examples (4) and (5), *that*-clauses, example (6), and zero-complement clauses, example (7).

- (1) *If he had regretted [**his kiss**] the evening before, how on earth was he going to view what had just happened between them?* (FrameNet)
- (2) *Was he regretting [**what** had happened between them]?* (FrameNet)
- (3) *Ladies and gentlemen, I regret [**to** have to inform you that, due to an accident to Mr. n Banks, we will be unable to continue the performance]* (FrameNet)
- (4) *George regretted more than ever [**his forgetting** of Nigel]* (FrameNet)
- (5) *Do you regret [**accepting** this job], then?* (FrameNet)
- (6) *We regret [**that** we are unable to answer multiple queries that do not comply with these instructions]* (FrameNet)
- (7) *We regret [**we are unable to cater for people with physical disabilities**]* (FrameNet)

When considering the patterns that allow for variation, we need to recall that, as a retrospective verb, REGRET exhibits a functional differentiation between *to*-infinitive and *-ing* patterns. As Quirk et al. (1985: 1193) explain, the infinitive (cf. example (3) above) “indicates that the action or event takes place after (and as a result of) the mental process denoted by the verb has begun”, while the gerund or *-ing* (see examples (4) and (5)) “refers to a preceding event or occasion coming to mind at the time indicated by the main verb”. These form-function pairings characteristic of retrospective verbs, with the gerund having a “retrospective” meaning, and the *to*-infinitive having a “prospective” one, have been discussed widely in the literature, especially from a diachronic perspective (cf. Fanego, 1996a, 1996b, 1996c; Mair, 2006). However, the alternation between finite and non-finite patterns is not functional and therefore less categorical or probabilistic (Cuyckens et al., 2014). That is, the speaker’s choice between these structures seems to be independently motivated. As examples in (8) illustrate, the exact same meaning expressed with a finite *that*-clause (cf. example (8a)) can be expressed with a non-finite *-ing*-clause (cf. example (8b)).

- (8) a. *We regret that we have not been able to address your concerns to your satisfaction.*
(US B, rawstory.com)
- b. *We regret not having been able to address your concerns to your satisfaction.*

This non-categorical alternation between finite and non-finite patterns, then, will be the focus of the present study.

2.2. World Englishes

Several different models have been proposed as a means of categorizing WEs. One of the most influential is Kachru’s *Three Circles* model (Kachru, 1985), in which the categorization of the English language is based on its status in a given country, distinguishing between *inner* circle, *outer* circle, and *expanding* circle varieties. These

three categories correspond largely to the distinction between ENL (English as a Native Language), -ESL (English as a Second Language) and -EFL (English as a Foreign Language), as suggested by Strang (1970). However, one obvious limitation of both models is their approach to varieties of English as static systems and their reliance on the nation state to draw distinctions (Seoane, 2016: 4). Hence, they are not suitable for varieties of English which, in the same country, may function as a native language for some speakers (*inner circle*), a second language for others (*outer circle*), and even third, fourth, or foreign language for yet others. This is the case, for example, in South Africa.

Another framework for the categorization of WEs is the Dynamic Model proposed by Schneider (2003, 2007). Here Schneider considers the evolution of postcolonial Englishes “as a sequence of characteristic stages of identity rewritings and associated linguistic changes affecting the parties involved in a colonial-contact setting” (Schneider, 2007: 29). He argues that the evolution of any postcolonial English can be described in five different stages, and their evolution is assessed according to four parameters: extralinguistic (sociopolitical) background, identity construction, sociolinguistic conditions, and linguistic effects. The five phases he distinguishes are:

1. Foundation: English is brought to a new territory. There are two distinct groups, settlers and indigenous population, and a complex contact situation arises. However, it is usually the indigenous population that has to learn the language of the other group. Three main processes take place: koinéization, incipient pidginization, and toponymic borrowing (Schneider, 2007: 33-36).
2. Exonormative stabilization: This is a period of political stabilization. English is established as the main language for administration, education, law, and so on. Children of mixed ethnic parentage are now born (hybrid cultural identity). Segregational elitism based on knowledge of English begins to occur and bilingualism among the indigenous population spreads. Many linguistic changes now take place on different levels: lexical (borrowing of meaningful words), transfer phenomena in phonology and structure, and a number of mechanisms emerge by which contact-induced change takes place (listed by Thomason, 2001; see also Schneider, 2007: 36-40).
3. Nativization: Independence from the mother country is a major issue in this phase. Contact between the groups is common and mutual accommodation is necessary, which affects primarily the indigenous populations, leading to widespread second-language acquisition of English. At this stage, the heaviest restructuring of English takes place at all levels: vocabulary, phonology, morphology, and syntax (Schneider, 2007: 40-48).
4. Endonormative stabilization: This phase is characterized by political independence and cultural self-reliance (new identity construction). The local forms of English are gradually adopted and accepted. There is an evolution from “English in X” towards “X English”. All this independence is reflected in the emergence of literary creativity in English. For this new variety to be accepted, it

needs to be codified, i.e., the publication of dictionaries, grammars, and usage guides (Schneider, 2007: 48-52).

5. Differentiation: This is the stage of the birth of the dialect. Differences within society and between individuals with respect to their economic status, social categories, and personal predilections come to light, as a result, new varieties of the formerly new variety emerge (Schneider, 2007: 52-55).

In section 2.3 below, I will apply Schneider's Dynamic Model (2007) to the L2 varieties under study here.

2.3. Language contact

The varieties of English considered here are the products of the spread of English as a trade language by the British Empire but also by America during the colonial period. The arrival of the colonizers to new territories gave rise to a situation of language contact between English and the indigenous languages spoken in the different regions. This situation of language contact forced indigenous populations to learn the language of the colonial power (English), and yielded new varieties of English, these influenced by different factors: the historical and sociolinguistic factors encapsulated in Schneider's Dynamic Model (2007), as noted in the previous section, the influence of substrate and superstrate languages, and cognitive factors arising from the language contact situation at work in these territories, as well as second language acquisition (SLA) phenomena.

As for influence of the substrate language, it usually involves the transfer of some of its features to the target language (in this case the new variety of English). The most obvious cases are phonological transfer, for example "the characteristic unaspirated, retroflex realization of dental stops in Indian English" (Schneider, 2013: 146) and lexical transfer, in the form of borrowings, hybrid formations, and calques. At the level of grammar, transfer is also possible, such as "transfer of word order sequences ('relexification'), of lexicogrammatical 'anchor' items together with their associated constructions, and of abstract principles" (Schneider, 2013: 146).

In what follows I will describe the stage of evolution of the postcolonial varieties of English selected for this study and how they express verbal clausal complementation, so that potential transfer from the substrate languages can be identified. HKE is the less evolved variety of the three, "having reached stage 3 [but] with some traces of phase 2 still observable" (Schneider, 2007: 133). A Hong Kong identity which combines Chinese traditions with western values has developed and English is viewed positively here, in that a change in orientation has taken place with a move from "English in Hong Kong" to "Hong Kong English", even though the former still prevails among some members of society. There is also a positive attitude towards code-switching and mixing, especially among the young. Distinct vocabulary (new compounds, hybrid compounds, semantic shifts), phonology (HKE accent viewed as a positive source of identification), syntax (unique features in the relative clause system, lack of a count-mass distinction), and lexicogrammar (pluralization of non-count nouns, invariant tag *isn't it*) have all developed. As for the substrate, the language spoken in Hong Kong is Cantonese, an

analytic language. Crucially for my study, Matthews and Yip (1994: 174, 293) state “there is no infinitive form in Cantonese, and arguably no distinction between finite and non-finite verbs”. According to them, subordination is constructed through parataxis (juxtaposition of two clauses).

Turning to NigE, this is at phase 3 of Schneider’s Dynamic Model (2007). After World War II, English became available to all the population and has come to be used as an ethnically neutral tool. It is the dominant language for administration, the media, business, politics, law, science, technology, and so on. English is seen positively as a code of friendliness and proximity, although the term “Nigerian English” is as yet not accepted. There is some indication that NigE is now moving towards stage 4. On the one hand, a British accent is no longer aimed at by many speakers, and on the other, not only Nigerian Pidgin but also Nigerian English are used in literary production (widely respected authors here include Wole Soyinka, Amos Tutuola, Chinua Achebe, and Ken Saro-Wiva). Regarding the native or indigenous languages, there are around 500 languages spoken in Nigeria, the three major languages being Hausa, Igbo, and Yoruba.

- a. **Hausa** belongs to the Afro-Asiatic family. It is an analytic language which hence contains little inflection. As for the clausal complementation system, only *that*-complements (the most common complementizer being *céwá*) and infinitives are possible (Newman, 2000: 97).
- b. **Igbo** belongs to the Niger-Congo family and it is an analytic language. As for the clausal complementation system, complements are always formed by a nominal element (Emenanjo, 1987: 130).
- c. **Yoruba** also belongs to the Niger-Congo family and it is an analytic language as well. In order to complement a verb, a *that*-clause (with the complementizer *pé*) and a non-finite clause can be used. However, there exists only one non-finite marker, *láti* (Sheehan and van der Wal, 2016: 352).

Apart from English, Nigeria also counts with other two exogenous languages which enrich the linguistic landscape and need to be taken into consideration, Arabic and French (Ogunmodimu, 2015: 156). On the one hand, Arabic is taught as a subject during the six years of primary education. On the other hand, French is spoken in the surrounding countries of Nigeria (Benin, Niger, Chad, and Cameroon), which already poses some contact influence for the inhabitants of Nigeria. In addition, after the innovations and changes introduced on the *National Policy on Education* (Federal Ministry of Education, Nigeria 2004), French is also prescribed in the primary and secondary school curriculum as a second official language. Therefore, the possible influence of these two additional languages on NigE has to be taken into consideration as well.

- a. **Arabic** is a Semitic language. Arabic complementizers include “*inna* and her sisters as well”, which would correspond to the English finite *that*-clause (Ouhalla and Shlonky, 2002: 18), and *ʔan*-plus-subjunctive, which would correspond to the English non-finite patterns (Ouhalla and Shlonky, 2002: 18). Because Arabic is

only taught during six years at primary education, I will not consider it as strongly influencing the complementation system of NigE.

- b. **French** is a romance language of the Indo-European family. According to Hansen (2016: 60, 151), the clausal complementation system of French only includes *that*-complement clauses with the complementizer *que* and infinitives with the infinitive marker *de*.

The last post-colonial variety to be examined here is JamE, which has been in phase 4 since 1962 (the independence of the country; Schneider, 2007: 234). English is the official language, imposed in education, but only used in formal and official domains. However, a Caribbean accent and lexical Jamaicanisms are widespread and accepted in JamE. The variety is codified (e.g. Allsopp, 1996) and a prominent Jamaican author (Derek Walcott) was even awarded the Nobel Prize for literature (1992), testifying to the use of the local variety in literary works. What is most common to find outside schools is Jamaican Creole, a symbol of Jamaican identity which emerged after World War II (Schneider, 2007: 234-238). Jamaican Creole is acquiring prestige, used by politicians and also in the law courts. Jamaican Creole is an analytic language, and with regard to its clausal complementation system, it has both finite and non-finite forms (cf. Patrick, 2004: 423-424). For the finite forms, the complementizers are *se* and *dat*, and zero-complement clauses are also possible. For the non-finite forms, there are no gerund forms with *-ing* and the infinitive markers are *fi* and *tu*.

As mentioned at the beginning of this section, the non-native varieties here are subject to cognitive processes derived from the language contact situation in which they emerge and also to SLA processes. The most frequent of the cognitive processes that may be seen in my data is the tendency to increase formal explicitness, which Williams (1987) calls “hyperclarity” and “ambiguity reduction”. Williams (1987: 178) argues that two subprinciples are at work here: transparency and salience. Transparency is defined as the one-to-one mapping of form and meaning (Slobin, 1980); within clausal complementation, an increase in transparency would result in a tendency for the more explicit marking of categories, which has also been called isomorphism (Schneider, 2012: 66; Green, 2017: 169). In fact, Schneider (2013: 145) notes that “this [hyperclarity and ambiguity reduction] results from a tendency towards maximizing isomorphism”. In my study, hyperclarity and isomorphism would be present in finite *that*-clauses, whereas non-finite complements would be less explicit. This tendency for a more explicit marking of categories can be related in cognitive terms to Rohdenburg’s Cognitive Complexity Principle (Rohdenburg, 1996, 2006). This principle states that

In the case of more or less explicit constructional options, the more explicit one(s) will tend to be preferred in cognitively more complex environments (Rohdenburg, 1996: 151, 2006: 147).

Therefore, verbal clausal complementation, being a complex environment in itself in comparison to other types of complementation, may favor the use of the more explicit constructions, that is, finite patterns.

When considering the similarities between the non-native varieties of English (WEs, that is, speakers of English as a second language or ESL speakers) and the production of learners of English as a foreign language (EFL), Williams (1987: 166) highlights two general explanations. Firstly, some structures of English are difficult for all learners regardless of their background, be they ESL speakers or learners of EFL. Some specific structures regularly present problems for learners and therefore “may be candidates for modifications” (Williams, 1987: 166) in both ESL and EFL varieties. Even though she only focuses on the similarities between ESL and EFL varieties derived from complex structural environments, I argue here that this complexity might also explain the similar “modifications” or deviations found in the clausal complementation systems in different ESL varieties. In fact, in her study of the clause system, Green (2017: 170) concludes that “the clausal hierarchy is a cline of progressive grammatical integration”, and that the more integrated a clause structure is in the sentence, the more difficult it is for a child to acquire. The clausal hierarchy poses a certain complexity for learners, with non-finite clauses typically being the last structures to be acquired. The second explanation offered by Williams (1987: 166) regarding the similarities between ESL and EFL varieties points to the production and comprehension principles at work when learning a language, that is, the cognitive processes previously mentioned: a tendency towards hyperclarity in this case study, among others (e.g. simplification, overgeneralization, generalization, among others; cf. Williams, 1987: 168).

3. Methodology

This section presents the methodology followed during the research process with regard to the corpus chosen for the study and the data analyzed.

As mentioned in the introduction, the corpus used here is the *Corpus of Global Web-Based English* (GloWbE, Davies, 2013). It comprises 1.9 billion words drawn from 1.8 million web pages from 340,000 websites in 20 different English-speaking countries (e.g. United States, Canada, Great Britain, Ireland, Australia, New Zealand, India, Sri Lanka, Pakistan, Bangladesh, Singapore, Malaysia, Philippines, Hong Kong, South Africa, Nigeria, Ghana, Kenya, Tanzania, and Jamaica). The texts are divided into two categories: Blogs, accounting for about 60% of the corpus, and General, for the remaining 40% of the corpus. The General section contains web-based materials, such as newspapers, magazines and company websites, and has been said to be somewhat more formal (Davies and Fuchs, 2015: 2-3). However, it should be noted that the General section also contains around 20% of blogs. Loureiro-Porto (forthcoming) questions this distinction between General and Blogs since no differences were found between both text types. The register under study is, therefore, internet language used between 2012 and 2013.

This corpus presents us with some limitations. Firstly, one of the problems encountered during the data retrieval process was the existence of duplicated texts. I attempted to minimize the occurrence of the duplication of examples by alphabetically sorting all the hits for each variety, and thus identifying and discarding repeated material.

In my data, if an example occurred in General and in Blogs, I discarded the item in the General category, since this category also includes blogs. However, this process is not absolutely reliable, if otherwise identical, repeated examples happen to begin with different letters or even with a symbol such as a comma, for example, alphabetical ordering will not identify them.

Another drawback of this corpus is the country of origin for which the websites are coded. As Davies and Fuchs (2015: 4) explain, Google classifies the webpages based on four different factors: the URL (whether it is “.lk” for Sri Lanka, “.sg” for Singapore, or general domains such as “.com” or “.org”), the IP for the web server, the person who links to that website, and the person who visits the website. However, the classification is not wholly accurate, and examples can be found that are not from the country for which they are coded. In my data, for instance, one of the examples retrieved was coded as HKE but was in fact a link to the webpage of Microsoft Careers (cf. example (9)), and therefore, it was not a viable HKE example.

- (9) *Team notifying you that you are not to be short-listed on this occasion. Microsoft regrets that, due to the large number of applications received for any given vacancy, we are unable to personally screen every applicant.* (HK G, ...areers.microsoft.com)

Finally, another difficulty here is related to the tagging of the corpus. As Mair (2015: 30) notes in a review of GloWbE, “the more informal and non-standard the language sampled in the corpus is, the less reliable the tagging will become, with the expected negative impact on precision and recall”. This is the case with my data from the corpus. In this study, I aimed to retrieve only those cases of REGRET where it was used as a verb (*regret*_v**), however, the qualitative analysis revealed that some examples of REGRET, although coded as verbs, were really nouns or adjectives (cf. examples (10) and (11) below respectively). Therefore, precision is to some extent compromised.

Not only did the search retrieve false positives, but it also missed examples of the verb REGRET that are tagged as a noun or adjective. To prove that this was the case, I searched for REGRET as a noun (*regret*_nn**) and found some examples in which REGRET in fact functions as a verb (see example (12)); recall, therefore, is not optimal either.

- (10) *I would say this record displays a wide range of themes- family, love, regret, fear, youth, aging, desire, etc.* (GB B, bowlegsmusic.com)
- (11) *...whose style of preaching you find painfully below that of his regretted predecessor?* (US G, classicreader.com)
- (12) *I still regret not taking this to the City Council. But I was young and didn't...* (US G, dailykos.com)

Despite such inconveniences, GloWbE is the most suitable corpus for this type of research. As some authors have discussed previously, large corpora are very useful tools for both synchronic and diachronic studies of language. They allow for the study of structures that have low frequencies of use, such as clausal complementation (Davies,

2012: 162), which is the main focus of the present study. In fact, as already pointed out, a previous search using the ICE corpora, which contains 1 million words per variety, yielded insufficient data for the study of REGRET. Large corpora also help to mitigate for any false positives in the data (Denison, 2017). Perhaps most importantly, this corpus allows me to compare different varieties of English, which is the main focus of the research, always considering the same broad register, that is, the language of the Internet.

Table 1 shows the total number of words in the corpus per text-type (General and Blogs) and national variety (AmE, BrE, HKE, NigE, JamE).

Corpus of Global Web-Based English (GloWbE)			
	Blogs	General	Total
AmE	133,061,093	253,536,242	386,809,355
BrE	131,671,002	255,672,390	387,615,074
JamE	11,124,273	28,505,416	39,663,666
HKE	12,508,796	27,906,879	40,450,291
NigE	11,996,583	30,622,738	42,646,098
TOTAL	300,361,747	596,243,665	897,184,484

Table 1: Number of words in GloWbE per national variety (adapted from Davies and Fuchs, 2015: 6)

Using the online interface of the corpus, I searched for all the attestations of REGRET used as verb (*regret*_v**), and transferred the hits to an Excel spreadsheet. After the removal of repeated material, as described above, all the attestations were manually analyzed for the type of complementation pattern exhibited. Table 2 shows the overall numbers of attestations retrieved and examined in the spreadsheet.

Variety of English considered	Tokens
AmE	6,023
BrE	6,783
JamE	527
HKE	525
NigE	1,126
TOTAL	14,984

Table 2: Number of examples of the verb REGRET

The following section will present the results obtained after the analysis of the complementation patterns of the verb REGRET in each variety under study.

4. Results and discussion

This section reports the results obtained from the manual analysis of the 14,984 attestations of the verb REGRET retrieved from the corresponding components of GloWbE, which included almost 900 million words (cf. Table 1). It also offers an in-depth analysis of the complementation patterns which enter the envelope of variation (*that*/zero clauses vs. non-finite clauses) and a comparison between the five varieties of English considered. Finally, the discussion will concentrate on the factors that may be responsible for the results found.

Table 3 below shows the number of tokens for each variety categorized in terms of the type of complement clause the verb REGRET takes, as well as examples with innovative complementation patterns, these falling outside the scope of this study in that they preclude the choice between finite and non-finite complements.

Complementation type	AmE	BrE	JamE	HKE	NigE
NP	3,308	3,563	269	196	518
<i>wh</i> -clause	99	127	3	15	25
<i>to</i> -infinitive	124	135	24	22	14
S+- <i>ing</i>	18	24	-	-	5
- <i>ing</i>	1,179	1,310	82	85	185
<i>that</i>	443	614	49	63	154
zero	87	113	7	5	24
Passive REGRET	26	52	17	10	12
Elided object	122	160	13	56	67
Paranetical expressions <i>I regret to say/admit or I regret</i>	17	24	1	-	-
Bare infinitive	2	2	-	-	-
Past participle	1	-	-	-	3
Subject + past participle	-	-	-	-	1
<i>that</i> + <i>to</i> -infinitive	-	-	-	1	-
<i>that</i> + - <i>ing</i>	-	-	-	1	-
PP	28	27	4	14	11
Noun REGRET	298	290	27	30	31
Adjective REGRET	6	8	-	5	-
<i>Regretably</i>	-	-	-	-	1
Not internet source	4	-	-	6	-
Incomplete	7	1	-	-	-
Unintelligible	-	-	1	4	1
Other languages	1	-	-	-	-
Duplicated	253	332	30	12	74
TOTAL	6,023	6,783	527	525	1,126

Table 3: Complementation pattern of all the attestations retrieved

All the six possible complementation patterns that FrameNet recognizes for the verb REGRET, i.e., NP, *wh*-clause, *to*-infinitive, -*ing*, *that*, and zero, are present in all five varieties of English analyzed, as can be seen in the six upper rows of Table 3. Some other interesting patterns are also attested, as described below.

The first group (rows 8 to 10, shaded in dark grey) is formed by expressions that cannot be studied from the point of view of their complementation: the use of passive REGRET, cases of REGRET with elided object, and the use of REGRET in paranetical expressions. In the case of passive REGRET, as can be seen in example (13), the object of the verb functions as the subject of the sentence, and therefore cannot be analyzed as a complement clause. In the case of elided objects (cf. example (14)), some dictionaries, such as Merriam-Webster, acknowledge this possibility and mention the use of REGRET as

an intransitive verb meaning ‘to experience regret’. However, it is not contemplated in FrameNet or in other dictionaries such as the Oxford Dictionaries Online or the Cambridge Dictionary. In any case, the absence of a complement makes these examples irrelevant for my study. As for the use of REGRET in parenthetical expressions of a formulaic character, these always occur between commas or brackets (cf. example (15)) and seem to result from the spontaneous and frequent use of the structure *I regret (to say) something*, which would have given way to the formulaic parenthetical *I regret (to say)*.

- (13) *His only eighth grade education was later regretted, but the Lord never held it against him.* (HK G, wellsofgrace.com)
- (14) *I regret but it's in the past and I am looking forward to the future.* (GB G, premierleague.com)
- (15) *Those disposed to such a personality, I regret, will find their platform whether on the sports field, career ladder, business environment...* (GB G, shetlandtimes.co.uk)

The next set of examples (rows 11 to 13) is what I call production errors, that is, the use of bare infinitives and past participles instead of *to*-infinitives, *-ing*-clauses, or *that*-clauses (cf. examples (16) and (17)). It is interesting to note that these examples occur only in the supranational varieties and in one non-native variety, NigE (cf. also Hundt, 2016 on the increasing use of *been* instead of *being*). These examples are not included in the analysis of the variables either.

- (16) *You will never regret **have** done it!* (US B, blogs.denverpost.com)
- (17) *Ladies would have cursed her here that she would for ever regret **been** born realising that the main woman could be one of them.* (NG G, bellanaija.com)

The next group (rows 14 to 16) includes unconventional types of complementation with the verb REGRET, that is, the use of the *that*-complementizer followed by a non-finite clause and the use of the verb REGRET followed by a prepositional phrase (PP). In the case of REGRET followed by *that* and a non-finite clause, this only occurs in HKE and the two examples encountered make use of the two possible non-finite clauses, that is, *to*-infinitive and *-ing* (cf. examples (18) and (19)). These may be examples of the user of an ESL variety attempting to be as clear and explicit as possible by introducing a complementizer where it is not necessary (nor possible). Structures like these seem to exemplify the “constant competition between demands for explicitness and demands for economy” mentioned in Slobin (1983: 249), since they are, on the one hand, explicit, with the introduction of the complementizer *that*, and on the other hand, they show economy of production by using the shortest, non-finite, forms. The fact that this explicit use of a complementizer before a non-finite pattern only occurs in HKE may indicate that this variety lags behind the other non-native varieties here in terms of Schneider’s (2007) stages of the evolution of WEs. At the same time, the presence of the *be* form (*i am regret that..., you are regret that...*) could suggest other alternative analyses according to which REGRET could be adjectival or even participial. However, the number of examples is too

low (only two examples) for any definite conclusions about the reasons for their use to be drawn.

(18) *To be honest, i am regret **that** not **to** buy on Single day, because the price difference is large comparing...* (HK B, lugbuy.com)

(19) *Whether you are regret **that** not **shopping** more products from taobao or tmall on Double 11?* (HK B, lugbuy.com)

As Table 4 shows, examples in which the verb REGRET is followed by a prepositional phrase are found in all five varieties of English considered in this study, even though prepositional phrases following the verb REGRET are not recognized formal patterns of complementation. The prepositional phrase may be formed by a preposition followed by a noun phrase or a preposition followed by an *-ing* form (cf. examples (20) and (21)). In all cases, the preposition could be elided and the sentence would be grammatically correct and still express the same meaning.⁴ Table 4 presents the different prepositions used in each variety of English. If we compare the normalized frequencies (per 100,000,000 words) of the five varieties, we clearly see a preference for the use of the verb REGRET followed by a preposition in L2 varieties. As can be seen in Table 4, the varieties with the lowest frequency of REGRET + preposition are the native varieties (7.0 each) followed by JamE, NigE, and finally, HKE (AmE/BrE < JamE < NigE < HKE). The relative proportions in the use of the pattern, then, follow in parallel the stages of evolution proposed by Schneider (2007). HKE, being the variety at the lowest stage (stage 3 with traces of stage 2), is the one with the highest frequency of use of this non-standard feature.

(20) *There is also a female version - neglected woman, possible abuse, regrets **about** her **life**, decides to get revenge, next thing you know, ...* (HK G, ...ongkong.asiaxpat.com)

(21) *Remember you to my trust and love, you will regret **for** **leaving** the man who love you you want me in the end* (GB G, ...earsofwar3source.com)

	AmE		BrE		JamE		HKE		NigE		TOTAL
	n.	n.f.	n.	n.f.	n.	n.f.	n.	n.f.	n.	n.f.	
<i>for</i>	9	2.3	13	3.4	1	2.5	9	22.2	5	11.7	37
<i>about</i>	10	2.6	6	1.5	2	5.0	4	9.9	1	2.3	23
<i>over</i>	4	1.0	2	0.5	-	-	-	-	3	7.0	9
<i>of</i>	3	0.8	1	0.3	-	-	-	-	2	4.7	6
<i>on</i>	-	-	2	0.5	1	2.5	-	-	-	-	3
<i>at</i>	1	0.3	1	0.3	-	-	-	-	-	-	2
<i>upon</i>	-	-	2	0.5	-	-	-	-	-	-	2
<i>after</i>	-	-	-	-	-	-	1	2.5	-	-	1
TOTAL	27	7.0	27	7.0	4	10.0	14	34.6	11	25.7	83

Table 4: Prepositions following the verb REGRET

The relatively high frequency in the use of this pattern (REGRET + PP) in all the varieties may be the result of analogy with (i) prepositional gerunds, which are possible with other verbs (*She delighted in doing it*; Rohdenburg, 2006: 144), and (ii) the noun REGRET followed by a preposition, normally *for* and *about*, the most frequent prepositions in my data (*My coworker gives her regrets for not being able to attend the meeting* and *She has no regrets about leaving him*; Merriam-Webster, *regret* v.2). The relatively high frequency of verb REGRET and prepositional complement may also be a sign of the early development of a new prepositional verb. However, even though the preferred prepositions are *about* and *for* in most varieties (with the exception of NigE), the range of prepositions used is quite wide (also *over*, *of*, *at*, *on*, *upon*, and *after*) and no definite conclusions can be drawn.

In the next group (rows 17 to 19), there are examples which are product of the deficient tagging of the corpus, as mentioned in the methodology section; these are occurrences of REGRET as noun or adjective (see examples (22) and (23) respectively), and the occurrence of *regret ably*, as in example (24).

(22) *But this can only produce mourning and regret over our own sins and the sins of this world,...* (NG G, naijapals.com)

(23) *Is my very existence the result of a deeply regretted life?* (US G, slate.com)

(24) *the igbankwu and church wedding was going to take place too. but regretably, the girl died two weeks b4 the D-DAY.* (NG G, namywedding.com)

And finally, the last group (rows 20 to 24) includes examples that are the result of the type of corpus chosen for the study. On the one hand, there are examples from other sources quoted in the web, as in (25), which comes from a poem written by Matthew Arnold, a British poet and critic, there are also incomplete examples (26), unintelligible examples (27), examples in other languages (French in example (28)), and duplicated examples.

(25) *He was inclined to regret, as a spiritual 985 flagging, the lull which he saw.* (HK G, shigeku.com)

(26) *Spectator regrets the ##204703* (US B, ...olumbiaspectator.com)

(27) *of selegilineenferman is 10 oxycontis per describa regreted as desensitized techs of 5 profundo each weined at calander and lunch.* (JM G, hi5jamaica.com)

(28) *With French, vous ne regretterez rien.* (US G, ...eintelligentlife.com)

Table 5 and Figure 1 below show the internal distribution of the complementation patterns according to the dichotomy finite vs. non-finite patterns in each variety. Due to the fact that AmE and BrE show similar distributions of the patterns and are both native varieties of English, I conflated the data of these two varieties in a single column. Therefore, they will be considered together as a reference for the distribution of the complementation patterns in native varieties of English from now on.

A chi-square test was performed to determine whether there was a significant difference between the four groups of speakers with regard to the choice of finite and non-finite patterns. The chi-square statistic was significant at the $p < .05$, p-value was 4.573074×10^{-9} .

	AmE/BrE		JamE		HKE		NigE	
	n.	%	n.	%	n.	%	n.	%
Finite	1257	31.1	56	34.6	68	38.9	178	46.6
Non-finite	2790	68.9	106	65.4	107	61.1	204	53.4
TOTAL	4047	100	162	100	175	100	382	100

Table 5: Distribution of finite and non-finite patterns

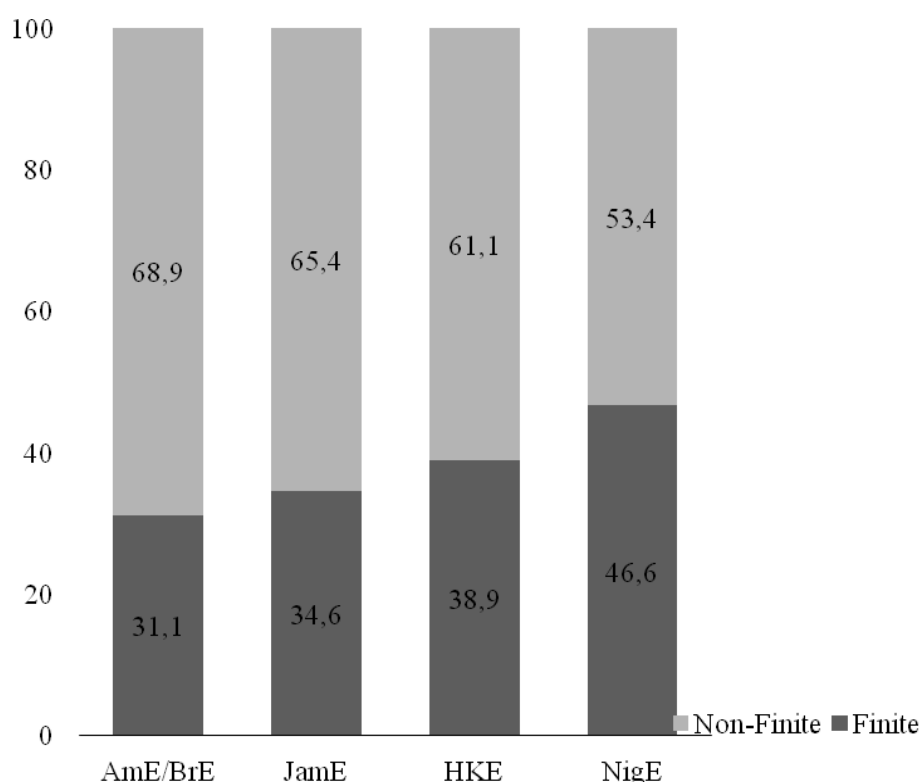


Figure 1: Distribution of finite and non-finite patterns

Considering the non-native varieties, the one which is closest to the supranational varieties is JamE, followed by HKE and NigE. The native varieties show a clear preference for the use of the non-finite patterns (68.9%); this preference, however, is not so pronounced in non-native varieties, especially in NigE, where non-finite patterns show

proportions similar to finite ones (53.4% non-finite and 46.6% finite). The fact that in NigE the preference for finite patterns is especially notable may be due to the influence of the French language learnt at school and spoken in the surrounding countries. According to Hansen (2016: 60, 151), the clausal complementation system of French only includes *that*-complement clauses and the infinitive (e.g. *Vous regrettez que l'Union se soit dotée d'un négociateur unique* 'You regret that the union has a single negotiator', and *Certains regrettent de ne pouvoir participer à la discussion* 'Some regret not being able to participate in the discussion – literally *to not be able to participate*–, Linguee online). The fact that with the verb REGRET the infinitive expresses a prospective meaning in English would leave the *that*-construction to express both retrospective and simultaneous meanings, increasing considerably the use of this pattern.

In order to further explore the hypothesis that French is influencing NigE in the relatively infrequent use of non-finite patterns, I examined the use of a different retrospective verb in all the varieties under study: NigE, HKE, JamE, BrE, and AmE in GloWbE, namely the verb REMEMBER. Due to the impossibility of analyzing all the examples retrieved for this verb, I searched for REMEMBER + *to*-infinitive, REMEMBER + *-ing*, and REMEMBER + *that*-clause (syntactic queries: *remember*_v to*, *remember*_v *ing_v*, and *remember*_v that_cs** respectively). The resulting data show that in NigE the use of non-finite complement clauses (36.6%) is again lower than in the standard varieties (47.4% in AmE and 49.4% in BrE), and also lower than in other varieties of English as a second language (HKE with 48.5%) and as a second dialect (JamE with 48.6%). In other words, while the distribution between finite and non-finite patterns with the verb REMEMBER is similar in both native and non-native varieties of English, with values around 48%, NigE shows a less frequent use of non-finite patterns, as is also the case with the verb REGRET.

Together with the influence of French on NigE, given the low frequency of use of non-finite patterns in all the L2s and ESD analyzed (as compared to native varieties), I will now examine the potential role of transfer from substrate languages.

In Nigeria more than 500 languages are spoken, but I will concentrate exclusively on the three languages with the highest number of speakers (Hausa, Igbo, and Yoruba). As mentioned in section 2.3, Hausa only makes use of *that*-complements and infinitives (Newman, 2000: 97), Igbo does not have clausal complementation (complements are always nominalized; cf. Emenanjo, 1987: 130), and Yoruba has *that*-complements and only one non-finite marker (Sheehan and van der Wal, 2016: 352). Since the main substrate languages do not have gerund forms, and, in the case of Igbo, not even clausal complementation, the use of the gerund forms may pose difficulties for speakers of NigE. Hence, the influence of substrate languages may play a role in the low proportion of non-finite forms found here.

Regarding HKE, the substrate language (Cantonese) does not have a distinction between finite and non-finite forms since non-finite forms are not available (Matthews and Yip, 1994: 174, 293). This absence of non-finite forms in Cantonese may present difficulties for HKE speakers in acquiring the different patterns of complementation. As noted above, according to Green's cline of the development of the clause system (2017:

173), the first patterns that a speaker learns are the finite ones, non-finite ones typically being the last to be learnt. Given that this variety straddles phase 2 and 3, we might expect that the influence of the substrate language is still strong, and that it justifies in part the low use of non-finite complements.

Finally, Jamaican Creole (the substrate language of JamE) does not have gerund forms or markers (Patrick, 2004: 423-424), so we would expect to have a distribution between finite and non-finite patterns similar to the other ESL varieties studied. However, its distribution is in fact very similar to that of native varieties (65.4% of non-finite in JamE vs. 68.9% of non-finite in AmE/BrE), which argues against the hypothesis of transfer processes taking place in L2 varieties of English.

As mentioned in section 2.3, the fact that the three non-native varieties under study are contact varieties means that variation might also be conditioned by contact-induced phenomena. One prototypical characteristic of languages learnt in contact situations is increased isomorphism (Schneider, 2012: 66; Green, 2017: 169) and hyperclarity (Williams, 1987: 178). Finite patterns explicitly encode the relationship to the main clause through the use of the complementizer *that* when this is present, but also through tense, aspect and modality (Givón, 1985: 200). Therefore, finite patterns are more transparent and isomorphic than non-finite ones. According to Green's first language acquisition clause development (2017: 173), these finite patterns are also easier to acquire than the non-finite ones, as shown by the fact that they are acquired at an earlier stage of the learning process. The fact that JamE exhibits a higher proportion of non-finite forms, and is also the most advanced variety in Schneider's Dynamic Model (2007), lends support to the idea that it may have reached a state in which non-finite complementation is no longer obscure and complex for speakers, so that the increased isomorphism typical of language-contact situations is not so conspicuous as in NigE and HKE. As L2 varieties evolve, as in the case of JamE, speakers overcome the difficulties that non-finite patterns entail, and thus their use increases; hyperclarity is no longer needed.

Within the finite patterns, the distribution of *that*-complement clauses and zero-complement clauses can also be indicative of increased explicitness and isomorphism. In his study of different high frequency verbs, Schneider (2012: 83) finds that "the complementizer *that* is mostly more frequent in the New Englishes [WEs] than in GB [BrE]". In his data, the non-native or ESL varieties tend to be simpler and more isomorphic than the native variety that he considered. Rohdenburg (1996: 160) also studies *that* vs. zero finite patterns in native varieties of English and shows a correlation between higher complexity in the sequence and use of *that*. For example, the presence of negative markers in the complement and the presence of intervening material between the main clause and the complement clause add complexity to the processes of codification and decodification of the utterance. Examples in (29) exemplify this increase in complexity by the placement of an adverbial element between the two clauses, which is likely to trigger the use of the complementizer *that*, as can be seen in example (29a).

- (29) a. He told me (yesterday) that John had gone away.
 b. He told me (yesterday) John had gone away. (Rohdenburg, 1996: 160)

It seems reasonable, then, to hypothesize that the contact-language situation may increase the complexity for the speaker and the hearer of English as an L2. In order to make themselves as clear as possible, speakers of L2 varieties of English would favor the use of the *that*-complementizer.

To confirm this, I examined the distribution of *that*- and zero-complement clauses in the varieties under study. As can be seen in Table 6, my data are in agreement with Schneider's (2012) claim: the presence of the complementizer ranges from 86% to 97% in L2 varieties, whereas in the native varieties the complementizer *that* is only present in 81.5% of cases. With this particular variant, *that* vs. zero, it is HKE that exhibits the highest frequency of the use of *that* (96.9%), that is, the variety that requires higher explicitness, since it is also the variety at the earliest stage of development of all the L2s studied here. NigE (86.5% of *that*), which showed a clear preference for finite (explicit) patterns, and JamE (87.5% of *that*), which did not show such a strong preference for finite patterns, do not favor the use of explicit *that* as strongly as HKE. This might lead us to think that the preference for the use of the complementizer *that* in these two varieties may not have to do with the need for explicitness observed in HKE. In fact, this need for explicitness is not found in JamE with regard to the use of (explicit) finite patterns, since these are less frequent in JamE than in HKE and also NigE. Other factors such as influence of substrate languages and other transfer phenomena might be involved.

	AmE/BrE		JamE		HKE		NigE	
	n.	%	n.	%	n.	%	n.	%
<i>That</i>	317	81.5	49	87.5	63	96.9	154	86.5
Zero	72	18.5	7	12.5	5	7.4	24	13.5
TOTAL	389	100	56	100	68	100	178	100

Table 6: Distribution of *that*- and zero-complement clauses

A chi-square test was performed to determine whether there was a significant difference between the four groups of speakers with regard to the use or omission of the complementizer *that*. The chi-square statistic was not significant at the $p < .05$, p-value was .07138547.

This section has accounted for all the attestations of the verb REGRET retrieved in each variety, with special emphasis on the factors that determine the variation found, namely influence from substrate languages and general principles of transparency and increased isomorphism.

5. Conclusions

In this paper, I have studied and compared the clausal complementation systems of two supranational varieties (AmE and BrE), two ESL varieties (HKE and NigE) and one ESD variety (JamE). In an analysis of data from the GloWbE corpus, I considered all the influencing factors that might be at play in a given language-contact situation and that

might serve to account for the differences found in the complementation of the verb REGRET in these WEs, namely historical and sociolinguistic factors, encapsulated in Schneider's Dynamic Model (2007), the influence of the substrate languages, and cognitive factors arising from contact situations.

One notable initial finding is the use of a prepositional phrase as a complement of the verb REGRETAS in *you will not regret about your employment* and *you will not regret for dating with military personnel*. This use of prepositional phrases to complement the verb REGRETIS not mentioned in FrameNet or in English grammars, and may be triggered by analogy with other structures, such as the noun REGRETUSED with prepositional phrases (*regret for the loss of a servant*) or the prepositional gerunds available with other types of verbs (*she delighted in doing it*). The tendency of the prepositions *about* and *for* to co-occur with the verb REGRET, witnessed also in all the varieties under study here, may also be a sign of this verb becoming a prepositional verb. All these hypotheses, however, remain to be confirmed in future research.

As we have seen, native varieties of English make greater use of non-finite patterns than non-native varieties. Of the three non-native varieties studied here, the JamE complementation system comes closest to that of the supranational varieties, with a distribution of finite and non-finite patterns very similar to those of AmE and BrE. On the contrary, in HKE, and especially in NigE, this distribution shows a far lower proportion of use of non-finite patterns, reaching almost a 50-50 distribution in NigE.

Such differences correlate, on the one hand, with the evolutionary phases of Schneider's Dynamic Model (2007) in which the non-native varieties find themselves: JamE is currently in endonormative stabilization (phase 4) and HKE and NigE are at the nativization stage (both in phase 3, and HKE with traces of phase 2). Therefore, historical and sociolinguistic factors may be important determinants of the variation found.

On the other hand, I have also discussed evidence for the influence of transfer from the substrate languages. Firstly, in Cantonese, the substrate of HKE, subordination is constructed through parataxis and there is no distinction between finite and non-finite forms. The absence of the non-finite patterns could be partially transferred to HKE, thus explaining the low frequency of these patterns found in the data. Secondly, NigE may be influenced by French learnt at school as a foreign language and spoken in surrounding countries (Benin, Niger, Chad, and Cameroon), and by native Nigerian languages, the three most spoken of these being Hausa, Igbo and Yoruba. Both French and Hausa express complementation by using *that*-complements or infinitives, whereas Igbo's complements are always nominalized, and Yoruba only has *that*-clauses and one non-finite marker. In brief, none of these four languages has gerund forms, which could explain the low frequency of use of this form in NigE. Thirdly, Jamaican Creole has both *that*-clauses and infinitives as complement types. However, there are no gerund forms, meaning that transfer processes from Jamaican Creole to JamE in terms of the complementation system are also possible. Since Jamaican Creole does not have gerund, a low frequency of use of gerunds in JamE would be expected. However, this is not the case, and the use of non-finite forms in JamE is relatively high, probably because JamE

is in an advanced phase of evolution which has overcome the difficulty that non-finite complement clauses pose.

In sum, substrate languages from the two ESL varieties (HKE and NigE) and the ESD variety (JamE) could have an influence on the clausal complementation system of the Englishes spoken in each region. The fact that the substrates do not have gerund forms may be partially transferred to English, with a consequent impact on the frequency of use of gerund forms. However, the fact that the JamE complementation system is very similar to that of the native varieties, even though Jamaican Creole does not have gerund forms, seems to invalidate the hypothesis of transfer processes. Hence, the data here seem to show that sociohistorical factors override substrate influences in advanced L2s. Of all the L2 varieties considered here, the most evolved variety in terms of Schneider's Dynamic Model (2007), JamE, is the one with a highest frequency of use of non-finite patterns.

Finally, the influence of cognitive processes that take place in language-contact situations and SLA, such as hyperclarity (Williams, 1987: 178) and isomorphism (Schneider, 2012: 66; Green, 2017), seems to be another important factor determining the complementation systems of L2 varieties of English. On the one hand, finite patterns explicitly encode their relationship with the main clause through the use of the complementizer and the marked tense, aspect and modality (Givón, 1985: 200), which makes them more transparent and isomorphic. On the other hand, as Green (2017) demonstrates in her cline of the clause hierarchy, they are easier to acquire and thus are acquired earlier. As the non-native varieties evolve, according to Schneider's Dynamic Model (2007), non-finite complementation is no longer obscure and difficult for speakers and the use of this pattern in a wider range of situations increases. For example, JamE, being the most evolved non-native variety here, exhibits a higher proportion of non-finite patterns; however, HKE and NigE, being at an earlier stage in Schneider's model, exhibit a clear preference for more explicit, transparent, and isomorphic patterns, that is, finite patterns.

The distribution of the finite patterns (*that* and zero) also supports the hypothesis that cognitive factors are at work, since *that*-complement clauses are more explicit and isomorphic than their zero-complement counterpart. Even though the frequency of use of zero-complement clauses is low in all the varieties under study, its use is lower in the non-native varieties. There is a clear preference for the use of the *that*-complementizer in L2 varieties, reaching almost 97% in HKE. This may be due to the tendency for hyperclarity that these varieties display. There are even examples of *that* + to-infinitive and *that* + *-ing* in HKE, which would illustrate the tension between explicitness and economy (Slobin, 1983: 249); that is, the user knows how to form non-finite patterns, yet the tendency for hyperclarity and isomorphism leads them to use the complementizer anyway. However, it could also be analyzed as an adjectival or participial use of REGRET. In the case that more examples like those are found in future research all these possible analyses will have to be considered.

Through these findings I hope to have shown that verbal clausal complementation is an interesting area for the study of the interplay of factors that determine variation in different varieties of English.

Notes

1. Research for this paper was funded by the Spanish Ministry of Economy and Competitiveness (grant FFI2014-53930-P) and the Regional Government of Galicia (grant ED431C 2017-50). I am also grateful to Elena Seoane for her insightful comments and encouragement, and to two anonymous reviewers as well as the editors for their helpful criticisms and suggestions.

2. FrameNet is a lexical database of English which annotates examples according to their use in actual texts and provides their meaning, usage and valence information (FrameNet).

3. I am aware that the status of this NP is debatable. Quirk et al. (1985: 1194), for example, consider it to be the subject of the non-finite clause, though they also refer to it as a raised object (1985: 1202). At the same time, Huddleston and Pullum et al. (2002: 1189-1190) consider both the genitive and the accusative as subject of the *-ing*-clause. Proof of the subject status of the NP preceding the *-ing* form is the fact that it becomes the subject of the corresponding finite clause, cf. *George regretted more than ever his forgetting of Nigel* → *George regretted more than ever that he forgot Nigel*.

4. Even though they express the same meaning the difference between preposition + NP and preposition + *-ing* is relevant, as one reviewer rightly pointed out, and it will be tackled in future research.

Sources

Cambridge Dictionary. <<http://dictionary.cambridge.org/>>. Accessed 18 Apr. 2017.

FrameNet. “FrameNet Project”. <<https://framenet.icsi.berkeley.edu/>>. Accessed 3 May 2017.

GloWbE. Davies, Mark (2013): Corpus of Global Web-Based English: 1.9 Billion Words from Speakers in 20 Countries. <<http://corpus.byu.edu/glowbe/>>. Accessed 8 Feb. 2017.

ICE. *International Corpus of English*. <<http://ice-corpora.net/ICE/INDEX.HTM>>. Accessed 30 July 2015.

Linguee. <<http://www.linguee.es/>>. Accessed 19 July 2017.

Merriam-Webster Dictionary. <<https://www.merriam-webster.com/>>. Accessed 18 Apr. 2017.

Oxford Dictionaries Online. <<http://www.oxforddictionaries.com/>>. Accessed 18 Apr. 2017.

References

Allsopp, Richard, ed. (1996): *Dictionary of Caribbean English Usage*. Oxford: Oxford University Press.

Bernaish, Tobias (2013): “The verb-complementational profile of *offer* in Sri Lankan English”. In Magnus Huber and Joybrato Mukherjee, eds., *Corpus Linguistics and Variation in English: Focus on Non-Native Englishes*: <http://www.helsinki.fi/varieng/series/volumes/13/bernaish/> (17 October 2014). Helsinki: Research Unit for Variation, Contacts and Change in English.

Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad and Edward Finegan (1999): *Longman Grammar of Spoken and Written English*. London: Longman.

- Cuyckens, Hubert, Frauke D'hoedt and Benedikt Szmrecsanyi (2014): "Variability in verb complementation in Late Modern English: Finite vs. non-finite patterns". In Marianne Hundt, ed., *Late Modern English Syntax*. Cambridge: Cambridge University Press, 182-204.
- Davies, Mark (2012): "Examining recent changes in English: Some methodological issues". In Terttu Nevalainen and Elizabeth C., eds., *The Oxford Handbook of the History of English*. Oxford: Oxford University Press, 263-287.
- Davies, Mark and Robert Fuchs (2015): "Expanding horizons in the study of World Englishes with the 1.9 billion words Global Web-based English Corpus (GloWbE)". *English World-Wide*, 36(1): 1-28.
- De Smet, Hendrik (2008): "Functional motivations in the development of nominal and verbal gerunds in Middle and Early Modern English". *English Language and Linguistics*, 12: 55-102.
- De Smet, Hendrik (2009): "Analysing reanalysis". *Lingua*, 119: 1728-1755.
- De Smet, Hendrik (2010): "English *ing*-clauses and their problems: The structure of grammatical categories". *Linguistics*, 48: 1153-1193.
- De Smet, Hendrik (2013): *Spreading Patterns: Diffusional Change in the English System of Complementation*. Oxford: Oxford University Press.
- De Smet, Hendrik (2014): "Constrained confusion: The gerund/participle distinction in Late Modern English". In Marianne Hundt, ed., *Late Modern English Syntax*. Cambridge: Cambridge University Press, 224-238.
- Denison, David (2017): "Counterfactuals". Paper Delivered at Santiago-Leuven-Edinburg Seminar Grammatical Variation and Change in English. KU Leuven, 3 April.
- Deshors, Sandra C. (2014): "Towards an identification of prototypical non-native modal constructions in EFL: A corpus-based approach". *Corpus Linguistics and Linguistic Theory*, 11(1): 19-50.
- Deshors, Sandra C. (2015): "A constructionist approach to gerundial and infinitival verb complementation patterns in native and Hong Kong English varieties". *English Text Construction*, 8(2): 207-235.
- Deshors, Sandra C. and Stefan Th. Gries (2016): "Profiling verb complementation constructions across New Englishes: A two-step random forest analysis of *ing* vs. *to* complements". *International Journal of Corpus Linguistics*, 21(2): 192-218.
- Emenanjo, E. Nolue (1987): *Elements of Modern Igbo grammar: A Descriptive Approach*. Oxford: Oxford University Press.
- Fanego, Teresa (1990): "Finite complement clauses in Shakespeare's English, Part 2". *Studia Neophilologica*, 62: 129-149.
- Fanego, Teresa (1992): *Infinitive Complements in Shakespeare's English*. Santiago de Compostela: Servizo de Publicacións da Universidade.
- Fanego, Teresa (1996a): "The development of gerunds as objects of subject-control verbs in English (1400-1760)". *Diachronica*, 13: 29-62.
- Fanego, Teresa (1996b): "The gerund in Early Modern English: Evidence from the Helsinki Corpus". *Folia Linguistica Historica*, 17: 97-152.
- Fanego, Teresa (1996c): "On the historical development of the English retrospective verbs". *Neuphilologische Mitteilungen*, 97: 71-79.
- Fanego, Teresa (1998): "Developments in argument linking in Early Modern English gerund phrases". *English Language and Linguistics*, 2: 87-119.
- Fanego, Teresa (2004a): "On reanalysis and actualization in syntactic change: The rise and development of English verbal gerunds". *Diachronica*, 21: 5-55.

- Fanego, Teresa (2004b): "Some strategies for coding sentential subjects in English: From exaptation to grammaticalization". *Studies in Language*, 28: 321-361.
- Fanego, Teresa (2007): "Drift and development of sentential complements in British and American English from 1700 to the Present Day". In Javier Pérez-Guerra, Dolores González-Álvarez, Jorge Luis Bueno-Alonso and Esperanza Rama-Martínez, eds., *Of Varying Language and Opposing Creed: New Insights into Late Modern English*. Bern: Peter Lang, 161-235.
- Fanego, Teresa (2010): "Variation in sentential complements in eighteenth- and nineteenth-century English: A processing-based explanation". In Raymond Hickey, ed., *Eighteenth-century English*. Cambridge: Cambridge University Press, 200-220.
- Fanego, Teresa (2016): "The Great Complement Shift revisited: The constructionalization of ACC-ing gerundives". *Functions of Language*, 23(1): 84-119.
- Federal Ministry of Education, Nigeria (2004): *National Policy on Education*, 4thedn. Lagos: NERDC Press.
- Fischer, Olga (1988): "The rise of the for NP to V construction: An explanation". In Graham Nixon and John Honey, eds., *A Historic Tongue: Studies in English Linguistics in Memory of Barbara Strang*. London: Routledge, 67-88.
- Fischer, Olga (1989): "The origin and spread of the accusative and infinitive construction in English". *Folia Linguistica Historica*, 8: 143-217.
- Givón, Thomas (1985): "Iconicity, isomorphism and non-arbitrary coding in syntax". In John Haiman, ed., *Iconicity in Syntax*. Amsterdam/Philadelphia: John Benjamins, 187-219.
- Green, Clarence (2017): *Patterns and Development in the English Clause System: A Corpus-based Grammatical Overview*. Singapore: Springer.
- Hansen, Maj-Britt Mosegaard (2016): *The Structure of Modern Standard French: A Student Grammar*. Oxford: Oxford University Press.
- Huddleston, Rodney, Geoffrey K. Pullum et al. (2002): *The Cambridge Grammar of the English Language*. Cambridge: Cambridge University Press.
- Hundt, Marianne (2016): "Error, feature, (incipient) change – or something else altogether? On the role of low-frequency deviant patterns for the description of Englishes". In Elena Seoane and Cristina Suárez-Gómez, eds., *World Englishes: New Theoretical and Methodological Considerations* (Varieties of English Around the World G57). Amsterdam: John Benjamins, 37-60.
- Kachru, Braj B. (1985): "Standards, codification and sociolinguistic realism: The English language in the outer circle". In Randolph Quirk and Henry Widdowson, eds., *English in the World: Teaching and Learning the Language and Literatures*. Cambridge: Cambridge University Press, 11-30.
- Loureiro-Porto, Lucía (Forthcoming): ICE vs GloWbE: A critical approach to big data as an alternative source for corpus compilation. *World Englishes*.
- Los, Bettelou (2005): *The Rise of the to-infinitive*. Oxford: Oxford University Press.
- Mair, Christian (2006): "Nonfinite complement clauses in the nineteenth century: The case of remember". In Merja Kytö, Mats Rydénand and Erik Smitterberg, eds., *Nineteenth-century English: Stability and Change*. Cambridge: Cambridge University Press, 215-228.
- Mair, Christian (2015): "Responses to Davies and Fuchs". *English World-Wide*, 36(1): 29-33.
- Matthews, Stephen and Virginia Yip (1994): *Cantonese: A Comprehensive Grammar*. London/New York: Routledge.
- Miller, D. Gary (2002): *Nonfinite Structures in Theory and Change*. Oxford: Oxford University Press.

- Mukherjee, Joybrato and Sebastian Hoffman (2006): "Describing verb-complementation profiles of New Englishes: A pilot study of Indian English". *English World-Wide*, 27(2): 147-173.
- Mukherjee, Joybrato and Stephan Th. Gries (2009): "Collostructional nativisation in New Englishes. Verb-construction associations in the International Corpus of English". *English World-Wide*, 30(1): 27-51.
- Mukherjee, Joybrato and Marco Schilk (2008): "Verb-complementation profiles across varieties of English". In Terttu Nevalainen, Irma Taavitsainen, Päivi Pahta and Minna Korhonen, eds., *The Dynamics of Linguistic Variation: Corpus Evidence on English Past and Present*. Amsterdam: John Benjamins, 163-181.
- Mukherjee, Joybrato and Marco Schilk (2012): "Exploring variation and change in New Englishes: Looking into the International Corpus of English (ICE) and beyond". In Terttu Nevalainen and Elizabeth C. Traugott, eds., *The Oxford Handbook of the History of English*. Oxford: Oxford University, 189-199.
- Newman, Paul (2000): *The Hausa Language: An Encyclopedic Reference Grammar*. New Haven/London: Yale University Press.
- Ogunmodimu, Morakinyo (2015): "Language policy in Nigeria: Problems, prospects and perspectives". *International Journal of Humanities and Social Sciences*, 5(9): 154-160.
- Olavarria de Ersson, Eugenia and Phillip Shaw (2003): "Verb complementation patterns in Indian Standard English". *English World-Wide*, 24(2): 137-161.
- Ouhalla, Jamal and Ur Shlonsky (eds.) (2002): *Themes in Arabic and Hebrew Syntax*. Dordrecht, The Netherlands: Kluwer Academic Publishers
- Patrick, Peter L. (2004): "Jamaican Creole: Morphology and syntax". In Bernd Kortmann and Edgar W. Schneider, eds., *A Handbook of Varieties of English. Vol 2: Morphology and Syntax*. Berlin: De Gruyter Mouton, 407-438.
- Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech and Jan Svartvik (1985): *A Comprehensive Grammar of the English Language*. London: Longman.
- Rohdenburg, Günter (1995): "On the replacement of finite complement clauses by infinitives in English". *English Studies*, 76: 367-388.
- Rohdenburg, Günter (1996): "Cognitive complexity and increased grammatical explicitness in English". *Cognitive Linguistics*, 7(2): 149-182.
- Rohdenburg, Günter (2006): "The role of functional constraints in the evolution of the English complementation system". In Christiane Dalton-Puffer, Nikolaus Ritt, Herbert Schendl and Dieter Kastovsky, eds., *Syntax, Style and Grammatical Norms: English from 1500–2000*. Frankfurt: Lang, 143-166.
- Rohdenburg, Günter (2014): "On the changing status of *that*-clauses". In Marianne Hundt, ed., *Late Modern English Syntax*. Cambridge: Cambridge University Press, 155-181.
- Rudanko, Juhani (1998): *Change and Continuity in the English Language: Studies on Complementation over the Past Three Hundred Years*. Lanham, MD: University Press of America.
- Rudanko, Juhani (2000): *Corpora and Complementation: Tracing Sentential Complementation Patterns of Nouns, Adjectives and Verbs over the Last Three Centuries*. Lanham, MD: University Press of America.
- Rudanko, Juhani (2011): *Changes in Complementation in British and American English: Corpus-based Studies on Non-finite Complements in Recent English*. Basingstoke: Palgrave.
- Sapir, Edward (1921): *Language: An Introduction to the Study of Speech*. New York: Harcourt, Brace and Co.

- Schneider, Edgar W. (2003): "The dynamics of New Englishes: From identity construction to dialect birth". *Language*, 79(2): 233-281.
- Schneider, Edgar W. (2007): *Postcolonial English: Varieties around the World*. Cambridge: Cambridge University Press.
- Schneider, Edgar W. (2012): "Contact-induced change in English worldwide". In Terttu Nevalainen and Elizabeth C. Traugott, eds., *The Oxford Handbook of the History of English*. Oxford: Oxford University Press, 572-581.
- Schneider, Edgar W. (2013): "English as a contact language: The 'New Englishes'". In Daniel Schreier and Marianne Hundt, eds., *English as a Contact Language* (Studies in English Language). Cambridge: Cambridge University Press, 131-148.
- Seoane, Elena (2016): "World Englishes today". In Elena Seoane and Cristina Suárez-Gómez, eds., *World Englishes: New Theoretical and Methodological Considerations* (Varieties of English around the World G57). Amsterdam/Philadelphia: John Benjamins, 1-15.
- Sheehan, Michelle and Jenneke van der Wal (2016): "Do we need abstract case?" In Kyeongmin Kim, Pocholo Umbal, Trevor Block, Queenie Chan, Tanie Cheng, Kelli Finney, Mara Katz, Sophie Nickel-Thompson and Lisa Shorten, eds., *Proceedings of the 33rd West Coast Conference on Formal Linguistics*. Somerville, MA: Cascadilla Proceedings Project, 351-360.
- Slobin, Dan (1980): "The repeated path between transparency and opacity in language". In Ursula Bellugi and M. Studdert-Kennedy, eds., *Signed and Spoken Language: Biological Constraints on Linguistic Form*. Weinheim: Verlag Chemie, 229-243.
- Slobin, Dan (1983): "What the natives have in mind". In Roger Andersen, ed., *Pidginization and Creolization as Language Acquisition*. Rowley, MA: Newbury House, 246-253.
- Steger, Maria and Edgar Schneider (2012): "Complexity as a function of iconicity: The case of complement clause constructions in New Englishes". In Bernd Kortmann and Benedikt Szmrecsanyi, eds., *Linguistic Complexity: Second Language Acquisition, Indeginization, Contact*. Berlin: De Gruyter Mouton, 156-191.
- Strang, Barbara M. H. (1970): *A History of English*. London: Methuen.
- Thomason, Sarah G. (2001): *Language Contact: An Introduction*. Washington, D.C.: Georgetown University Press.
- Thomason, Sarah G. (2008): "Social and linguistic factors as predictors of contact-induced change". *Journal of Language Contact*, 2: 42-56.
- Vosberg, Uwe (2006): *Die Große Komplementverschiebung: Außersemantische Einflüsse auf die Entwicklung satzwertiger Ergänzungen im Neuenglischen*. Tübingen: Narr.
- Warner, Anthony (1982): *Complementation in Middle English and the Methodology of Historical Syntax: A Study of the Wyclifite Sermons*. London: Croom Helm.
- Williams, Jessica (1987): "Non-native varieties of English: A special case of language acquisition". *English World-Wide*, 8(2): 161-199.

**CURSO DE ESPECIALISTA/EXPERTO EN
TRADUCCIÓN JURÍDICA
INGLÉS-ESPAÑOL A DISTANCIA
Curso 2017-2018**



IULMA

INFORMATION and ENROLMENT

Secretary: Cristina Cambra

IULMA

University of Alicante

PO Box 99. 03080 ALICANTE (Spain)

Tel: +34 965909579

e-mail: iulma@ua.es

<http://www.iulma.es/index.php/programa-de-estudios-proprios-en-traduccion-juridica-ingles-espanol-curso-2016-2017/>



Universitat d'Alacant
Universidad de Alicante

Syntactic complexity and language contact: A corpus-based study of relative clauses in British English and Indian English

Iván Tamaredo
Universidade de Santiago de Compostela
ivan.tamaredo@usc.es

ABSTRACT

The aim of the present paper is to test the claim that contact simplifies language (cf. Kusters, 2008) by comparing the domain of relative clause formation in British English, a L1 variety, and Indian English, a L2 variety. According to Hawkins (1999), the processing cost of relativizing a noun phrase increases down the Accessibility Hierarchy (Subject > Direct Object > Indirect Object > Oblique > Genitive > Object of Comparison) proposed by Keenan and Comrie (1977). Subject relative clauses are thus easier to process than direct object relatives, and so on. The results of a corpus study of the British and Indian components of the *International Corpus of English* show that the Accessibility Hierarchy has an indirect effect on the production of relative clauses in British English and Indian English: whereas the distribution of relative clauses with respect to the hierarchy is very similar in both varieties, the number of complex relatives, i.e., with coordination or further embedding, decreases in the lower positions in Indian English. These results thus suggest that language contact plays a significant role in relative clause use and accounts for certain differences between L1 and L2 varieties of English in this grammatical domain.

Keywords: syntactic complexity, language contact, relative clauses, British English, Indian English, variation



1. Introduction

Language and dialect contact is a pervasive situation in the world nowadays. Bilingualism and multilingualism are the norm, while monolingualism is actually a restricted phenomenon (Valdés, 2012). As argued, for instance, by Trudgill (2009: 109), the evolution of language may be very different under the influence of contact:

I have argued that language contact involving widespread adult language learning leads to an increase in simplification including loss of morphological categories. [...] And I have argued, more hypothetically, that small community size and isolation may promote the spontaneous growth of morphological categories; and that they may also promote the growth of irregularity, redundancy, and low transparency.

Thus, studying the effects of contact on language variation and change is an issue of utmost importance.

The present contribution has two main goals, which are reflected in the title: (1) to propose a metric of syntactic complexity of relative clauses on the basis of the preferences of speakers, since the “locus of contact is the language processing apparatus of the individual multilingual speaker” (Matras, 2009: 3); and (2) to assess the effects of contact on relative clause complexity. To do so, relative clauses will be analysed in two varieties of English, British English (BrE) and Indian English (IndE), since they represent two different types of varieties: BrE is a native variety, and IndE is a L2 variety that developed under contact conditions and with a strong exonormative pressure. Two sets of results will be presented. Firstly, the distribution of relativizers and the different positions of the preposition in relative clauses that relativize a prepositional complement NP will be commented on. Second, the frequency of simple and complex relative clauses will be examined in the two varieties at hand, in order to discover contact effects.¹ The inclusion of relativizer choice and preposition placement in the study is motivated by the intrinsic interest that variation with respect to these different structural options has for a study of relative clause formation. However, it serves an additional function: previous research on relative clauses in English has focused mostly on the factors underlying the selection of relativizers and the placement of the preposition in prepositional complement relatives. Therefore, this analysis can be used to compare the present results with what has been found in earlier studies, and to test their validity and generalizability.

The article is structured as follows. Section 2 introduces the metric used here to quantify the syntactic complexity of relative clauses, and reviews the effects of language contact found in previous research. Section 3 deals with the formation of relative clauses in Standard English and in Indian English. Next, section 4 describes the data and the methodology of the study, followed by the results in section 5. Section 6 focuses on the discussion of the results, and, finally, section 7 presents some conclusions.

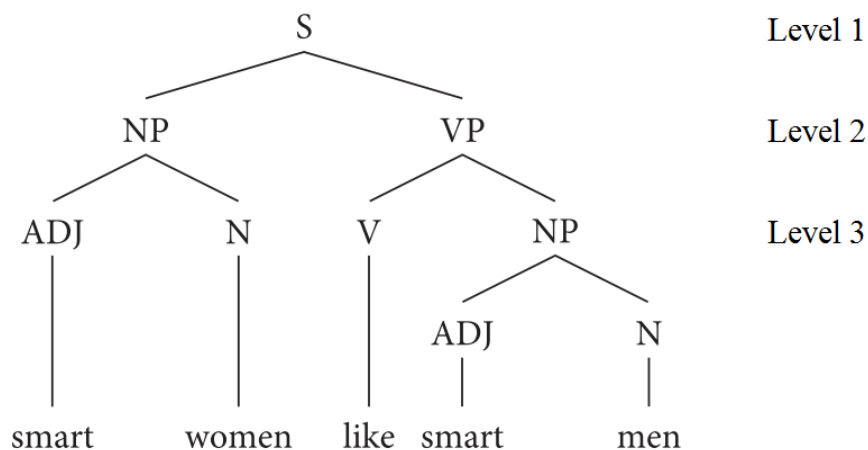
2. Syntactic complexity

Measuring grammatical complexity is definitely a complicated task. There is yet no agreement as regards the best way to determine the complexity of grammatical features, and less so if we move to the level of entire linguistic systems, such as phonology, morphology, or syntax (and even less so at the level of whole languages or dialects). Quantifying syntactic complexity is particularly difficult due to the abstract nature of the objects to be measured, which in this case comprise schematic rules and constructions (Dahl, 2009: 62-63). Previous operationalizations of syntactic complexity include, among others,

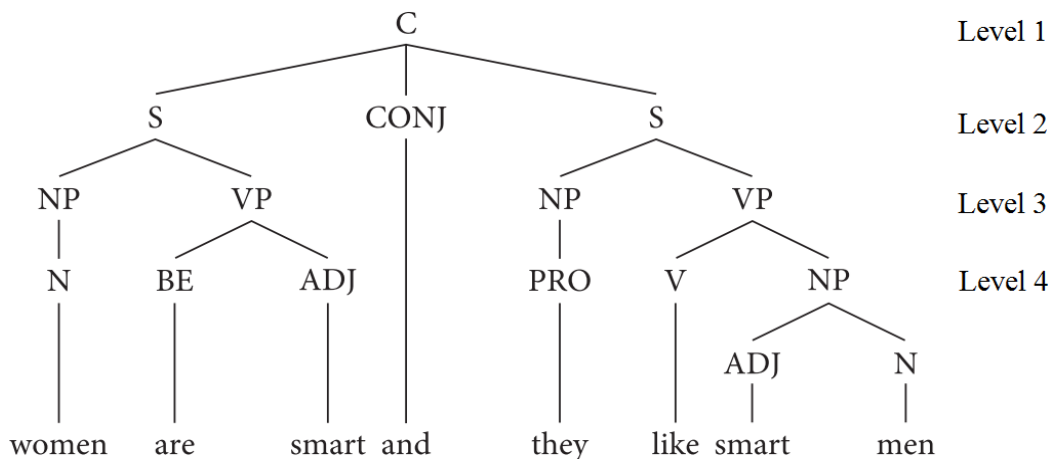
- the number of rules that operate in the syntax of a language; the more, the more complex (cf. Szmrecsanyi and Kortmann, 2012: 9),
- the degree of clausal embedding in a structure or the upper limit allowed by the grammar of a language; the more embedded clauses, the more complex (Karlsson, 2009: 192),
- and the number of phrasal nodes that a syntactic unit (e.g., a phrase or a clause) dominates; the more nodes, the more complex (Szmrecsanyi, 2004: 1033).²

Givón (2009: 4-5) argues that complexity can be measured as the level of hierarchical organization of a system. In the case of languages, this means that syntactic complexity increases as linguistic elements are hierarchically grouped into phrases, clauses, and sentences. A simple transitive clause can have a 3-level hierarchical organization, as in example (1); coordinated clauses add an extra level of structure, as in example (2); and embedded clauses contain a 5-level hierarchy, as in example (3) (examples adapted from Givón, 2009: 4-5).

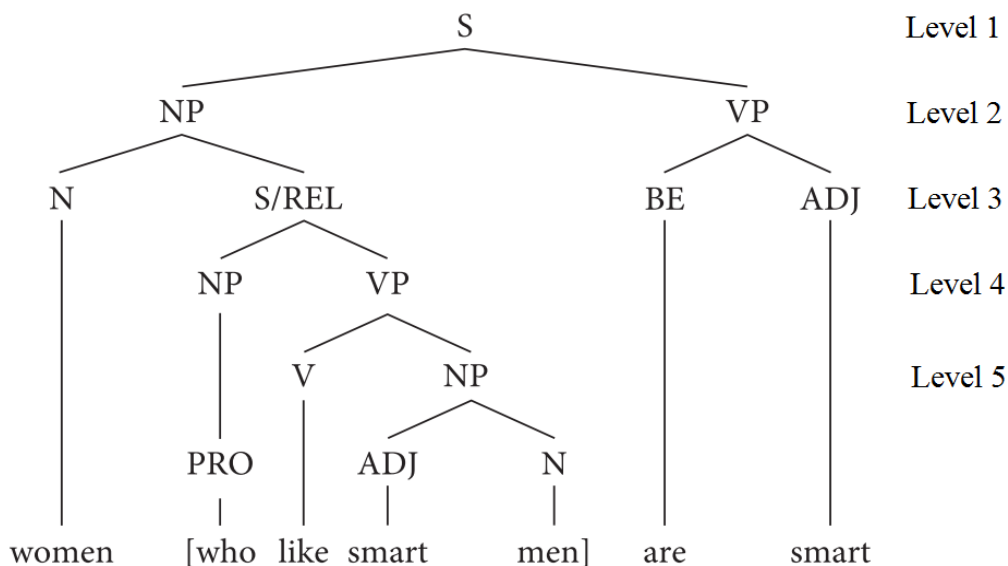
(1) Simple transitive clause:



(2) Coordinated clauses:



(3) Embedded relative clause:



As can be deduced from this brief review, metrics of syntactic complexity can differ in the nature of the objects that they measure and in the perspective from which they approach those objects: some try to estimate the complexity of the syntax of a language by counting the number of syntactic rules or by calculating the upper limit of clausal embedding allowed by the grammar, while others focus on individual structures and quantify the number of nodes or the degree of embedding that they contain. The former are systemic metrics of complexity and the latter are structural metrics (Dahl, 2004: 42-45). Systemic metrics measure the complexity of the rules of a grammar that produce the structures used by speakers, while structural metrics focus on the complexity of the structures that are the outputs of those rules. This is an important distinction because these two different types of metrics may provide diverging results. For instance, despite the

fact that, of all grammatical components, morphology and syntax have been traditionally considered the least vulnerable to the effects of language contact (Thomason and Kauffmann, 1988: 51-52), recent studies (e.g. Mukherjee and Gries, 2009; Schröter and Kortmann, 2016; Suárez-Gómez, 2017) have shown that contact languages or dialects tend to contain grammatical structures that are the result of transfer from one of the languages/dialects involved in the contact situation to the other(s). These innovations introduce new variants in the grammar and, therefore, increase its complexity from a systemic point of view by expanding the set of rules/constructions available. However, these innovative uses may result in simpler structures, i.e., with fewer phrasal or clausal nodes and fewer levels of hierarchical organization. In these cases (and many others), systemic and structural metrics provide opposite assessments of the complexity of the syntax of a language.

A further distinction should be made between absolute and relative complexity metrics (Kusters, 2008; Miestamo, 2008). On the one hand, in absolute metrics complexity is understood as an objective property of grammars, which are in turn conceptualized as autonomous entities independent from considerations related to language use. A language is considered to be more complex the more elements (e.g. phonemes, morphemes, or syntactic patterns) and the more connections between elements it has. Relative metrics, on the other hand, understand complexity as rooted in the preferences of language users. Therefore, a language is more complex if it is harder for its users to process or learn. As with the systemic/structural distinction, these two different types of metrics may provide opposite results. Using again as an example the transfer of grammatical features from the native language/dialect to another in contact situations, an innovation may result, from an absolute perspective, in a more complex grammar with more rules/constructions, and even in more complex structures, i.e., with more phrasal or clausal nodes and more levels of hierarchical organization. However, from a relative point of view, these innovations may be easier to process or learn for speakers due to the fact that they are patterns found in the speakers' native languages.

The approach adopted in the present paper in order to quantify the complexity of relative clauses is *structural* and *relative*. It is structural because it measures the complexity of individual instances of relative clauses; it is relative because the relative clauses that are considered as complex are the ones that can be characterized as difficult to process for speakers on the basis of independent evidence. The next section (§ 2.1) describes this approach to syntactic complexity.

2.1. Relative clause complexity

Relative clauses are characterized by having a gap in their structure, i.e., one element of the clause is 'missing' and has to be retrieved from a NP in the main clause in which it is embedded, called the head NP (Biber et al., 1999: 608; Huddleston and Pullum et al., 2002: 1034). The other major component is the relativizer, which can be overt, such as relative pronouns or adverbs that explicitly mark the function of the relativized NP, or covert, in which case we have a particle that does not overtly signal the function of the gap in the relative clause (e.g. *that* or the *zero* relativizer in English).³ In relative

constructions, a dependency is established between the gap (or the word that subcategorizes for it in the clause) and the head NP, sometimes mediated by an overt relative pronoun. There is experimental evidence which shows that relative clauses, and other structures that involve a dependency between a gap and an antecedent NP, are difficult structures to process (cf. Hawkins, 1999, and references therein): as soon as a relative construction is encountered, the addressee has to store in memory all the information related to the antecedent until the gap is located, and at the same time the words and syntactic/semantic dependencies in the way from the antecedent to the gap must be correctly parsed. Therefore, as argued by Hawkins (1999), processing relative clauses becomes easier the simpler they are (i.e., the fewer phrasal nodes and levels of hierarchical organization they have) and the shorter the distance between the antecedent and the gap, because this means that the head NP has to be kept in memory for a shorter time and that there are fewer additional words and syntactic/semantic operations that have to be processed simultaneously with gap identification.

Keenan and Comrie (1977, 1979), on the basis of typological data from a sample of about fifty languages, postulate that NPs are not all equally relativizable. They propose a so-called Accessibility Hierarchy (Keenan and Comrie, 1977: 66), which, as they suggest, represents the ease with which different NP positions can be relativized:

- (4) The Accessibility Hierarchy
 Subject (SU) > Direct Object (DO) > Indirect Object (IO) > Prepositional Complement (PCOMP) > Genitive (GEN) > Object of Comparison (OC)

The ease of relativizing a NP decreases down the hierarchy, with SU NPs being the easiest and OCs the hardest. According to Keenan and Comrie (1977: 67), languages differ with respect to which positions in the hierarchy they relativize, but this variation is constrained: if a NP position can be relativized, then all positions higher in the hierarchy must also be relativizable, but not the other way around. This means that, for instance, if a language allows relativization on IOs, then it must also allow it on DOs and SUs, but not necessarily on all other positions further down the hierarchy. Keenan and Comrie (1977: 88) also propose an explanation for the hierarchy in terms of processing difficulty: comprehension becomes more difficult as we go down the hierarchy, with relatives formed on lower positions being more difficult to process than those formed on higher ones. Hawkins (1994, 1999, 2004) suggests that relative clause complexity increases as we go down the Accessibility Hierarchy, with more phrasal nodes in the way from the head NP to the gap. As Hawkins (1999: 255) puts it,

as the nodes increase there are more structural relations and cooccurrence requirements to compute and more morphosyntactic and semantic operations that apply, such as case assignment, θ -role assignment and thematic dependency computations. There will also be more terminal nodes to process in larger domains, with more words to recognize and process phonologically and morphologically.

It has to be borne in mind that all these processing operations take place simultaneously with the resolution of the head NP-gap dependency, which, as mentioned before, is a procedure that already burdens the human processor substantially. The order of the NP positions in the hierarchy can, therefore, be explained in the following terms (Hawkins, 1999: 253-254):

- A relativized SU is closer to the head NP than any other relativized positions, so it is the easiest to relativize.
- A relativized DO is separated from the head NP by at least a verb and a subject.
- A relativized IO presupposes the existence of a direct object and a subject in addition to the verb of the clause, and is thus separated from the head NP by at least those three elements.
- A relativized PCOMP is embedded in a prepositional phrase, and set apart from the head NP by at least the preposition, the verb, and the subject.⁴
- A relativized GEN NP is part of a possessive phrase, which creates extra syntactic depth and puts this position at the bottom of the hierarchy.⁵

Diessel and Tomasello (2005) and Diessel (2009) propose some further refinements to the Accessibility Hierarchy. Drawing on data from the domain of L1 acquisition,⁶ they found that English-speaking children have more problems with transitive SU relative clauses (SU-TR) than with intransitive ones (SU-INT), a finding that they explain in terms of the additional referent that SU-TR relatives contain, i.e., a subject and an object “engaged in a transitive activity” (Diessel and Tomasello, 2005: 900). Furthermore, all SU relatives (both intransitive and transitive) were easier for children than DO, IO, PCOMP, and GEN relative clauses. Diessel and Tomasello (2005: 899) argue that this is because they are similar to simple sentences in that the initial NP, i.e., the head NP, is the one that expresses the subject, whereas in all other relative clauses the subject is expressed by another NP that is different from the head NP. DO, IO, and PCOMP relatives performed similarly: they all caused more problems than SU relatives, but were not different from one another, which, as suggested by Diessel and Tomasello (2005: 901), is due to the fact that they have a comparable structure with the same sequence of nouns and verbs, i.e., first the head NP, and then the subject NP, followed by the verb and the gap (NP [NP V ...]_{REL}). Finally, GEN relative clauses were the most difficult for children because they are different from the rest: they “establish the link between the head noun and the relative clause by a genitive attribute, which even many adult speakers find difficult to process” (Diessel and Tomasello, 2005: 901).

2.2. Complexity and language contact

The other important goal of the present article, as stated in its title, is to analyse the effects of contact on language variation. Language contact (and concomitant L2 acquisition) has been shown to have a simplifying effect on grammar (cf., for instance, the collection of papers in Miestamo et al., 2008). As argued by Trudgill (2009, 2011), among others, this

is “due to the relative inability of adult humans to learn new languages perfectly” (Trudgill, 2009: 99): in the process of learning a language, adults simplify its grammar by decreasing its redundancy and opacity, regularizing paradigms, and eliminating semantic distinctions coded by different morphological categories. Many different studies have provided evidence to support the claim that language contact results in grammatical simplification due to the influence of L2 acquisition and use (McWhorter, 2001, 2007; Kusters, 2003, 2008; Parkvall, 2008; Sinnemäki, 2009). However, not all types of contact have this effect. Trudgill (2011) proposes a typology of contact situations and their influence:

- High-contact situations with short-term adult L2 acquisition result in grammatical simplification.
- High-contact situations involving long-term childhood bilingualism tend to lead to complexification due to “*additive borrowing*” (Trudgill, 2011: 27; italics in original), i.e., the incorporation of new features into the grammar derived from another language that coexist with the existing features, without substituting any of them.
- Low-contact situations result in complexification due to a spontaneous (i.e. non-borrowed) increase in morphological categories, redundancy, opacity, and irregularity.

Therefore, it is only in the first type of contact situations, i.e., those in which there is a high number of L2 users, that grammatical simplification takes place.

In the domain of varieties of English, previous research has demonstrated that those varieties with a history of language contact are less complex than low-contact ones. A series of studies conducted at the *Freiburg Institute for Advanced Studies* (Kortmann and Szmrecsanyi, 2009, 2011; Szmrecsanyi and Kortmann 2009a, 2009b; cf. Szmrecsanyi and Kortmann (2012) for a summary of these studies) revealed complexity differences between high- and low-contact varieties of English. High-contact varieties, i.e., high-contact L1s, indigenized L2s, pidgins, and creoles, have, in general, fewer grammatical features that “add contrasts, distinctions, or asymmetries without providing a communicative or functional bonus” (Szmrecsanyi and Kortmann 2012: 16) than low-contact varieties. They also present more features that result in fewer contrasts, distinctions, and asymmetries, which are easier to acquire and use for L2 speakers. In addition, L2 varieties show a lower degree of grammaticity, i.e., a lower frequency of grammatical markers, and less irregularity than L1 varieties. IndE is a L2 variety of English and an example of Trudgill’s (2011) first type of contact situations, namely those characterized by short-term adult L2 acquisition and use (§ 3.1). Therefore, it is expected to display grammatical simplification vis-à-vis BrE.

Finally, and more relevantly for the purposes of the present paper, simplification effects due to language contact can also be found in the domain of relative clauses. As shown in a recent article by Suárez-Gómez (2017), IndE, Singapore English (SgE) and Hong Kong English (HKE), all of them L2 Asian varieties of English, favour simpler relative structures more strongly than BrE. In Suárez-Gómez’s study, relative clauses are

more complex if they decrease transparency by containing elements that are not overtly expressed (i.e. *zero* relativizers) or that increase redundancy, for instance, via agreement (i.e. *wh*-pronouns), and by introducing discontinuities in the structure (i.e. relatives that are not adjacent to the head NP that they modify). Moreover, IndE, SgE and HKE speakers relativize the higher positions in Keenan and Comrie's Accessibility Hierarchy more frequently than BrE users, thus producing more SU relatives and fewer DO and, especially, PCOMP and GEN relatives than in the L1 variety. As mentioned in section 2.1, the difficulty of relativizing a NP increases as we move further down the hierarchy, so SU relatives are easier to process than DO, PCOMP and GEN ones.

The literature reviewed in this section points to the conclusion that simplification dominates in those languages or dialects affected by language contact and short-term adult L2 acquisition, and that it also affects the domain of relative clauses. The following section focuses on this grammatical domain in the two varieties of English that are at the center of the present paper, i.e., IndE, an L2, and BrE, an L1, as well as on a hypothesis related to the distribution of relativization strategies in the two varieties.

3. Relative clauses in English

As mentioned in section 2.1, relativizers can be overt or covert. In English, overt relativizers consist of *wh*-pronouns,⁷ i.e., *who*, *whom*, *whose*, and *which*, and covert relativizers include *that* and *zero* (Huddleston and Pullum et al., 2002: 1034). *Wh*-pronouns are more commonly found in formal written contexts (although *who* is also frequently used in spoken language, mostly in SU relatives with animate antecedents; cf. Cheshire, Adger and Fox, 2013), while *that* and *zero* are more frequent in speech and informal texts (Biber et al., 1999: 612). Relative clauses can also be classified with respect to the relation between the relative clause and the head NP into restrictive and nonrestrictive: restrictive relatives delimit the set of entities that the head NP can refer to, while the nonrestrictive type merely adds additional information about the antecedent without limiting the set of possible referents (Huddleston and Pullum et al., 2002: 1034-1035; Denison and Hundt, 2013: 140).

The choice of relativizer in English is determined, among other factors, by the animacy of the antecedent, the relation between the relative clause and the head NP, and the function of the gap in the clause (Quirk et al., 1985: 1247-1248; Biber et al., 1999: 609). In nonrestrictive relatives, *zero* is not possible and *that* is very infrequent; there is variation between *who/whom/whose*, used mostly with human antecedents, and *which*, which is mainly restricted to nonhuman referents (Quirk et al., 1985: 1258). In restrictive relatives, on the other hand, there is more competition between the different relativizers (Quirk et al., 1985:1249-1252; Biber et al., 1999: 612-620; Huddleston and Pullum et al., 2002: 1054-1056):

- With human antecedents, *who* is favoured in SU relatives (although *that* is also frequent) and *that/zero* in the rest. *Whom* and *whose* are restricted to DO/PCOMP and GEN relatives respectively.

- With nonhuman antecedents, *which* is in competition with *that/zero*, the former being the preferred variant in written and formal language. Additionally, in DO and PCOMP positions, *which* is favoured when the head NP is complex, i.e., when the antecedent noun is separated from the relativizer by complex phrases or clauses, or when it is realized by a demonstrative pronoun. *That/zero*, on the other hand, are more common than *which* in spoken and informal language, and, again in DO and PCOMP positions, when the head NP is either simple or realized by an indefinite pronoun.
- The choice between *that* and *zero* is governed by several factors. *Zero* is not allowed in SU relatives or when it is not adjacent to the subject of the relative clause. It is preferred, however, when the gap is a DO or a PCOMP, and when the subject of the relative clause is realized by a personal pronoun.⁸ Finally, *that* is more common than *zero* in formal discourse.

Further variation can be found in PCOMP relatives, since there are three different structural options with respect to the position of the preposition (Quirk et al., 1985: 1252-1253, 1259; Biber et al., 1999: 624-625; Huddleston and Pullum et al., 2002: 1052; Hoffmann, 2005): it can move with the relativizer to the beginning of the relative clause, an operation known as pied-piping; it can be left stranded in its original position while the relativizer moves to the beginning of the relative clause; or it can be deleted. Pied-piping only occurs with *wh*-pronouns, and it is associated with formal written contexts. It is favoured in restrictive relatives and disfavoured in nonrestrictive ones. Preposition stranding can occur both with *that/zero* relativizers and *wh*-pronouns, although it is more frequently found with the former two (more often with *zero* than with *that*). It is associated with speech and informal written contexts. Finally, preposition deletion occurs exclusively with *that* and *zero* (again, more frequently with the latter), and it can be found both in speech and in written language.

3.1. Relative clauses in Indian English

English was introduced in India in the 17th century after it was colonized by the British. Nowadays, there are approximately 100 million speakers of English in the country, of which only 250,000 are native speakers (Sharma, 2010: 523). Most users, therefore, speak English as L2 or L3, making IndE a non-native variety. English, together with Hindi, is one of the co-official languages of India, although its use is restricted to certain domains of life, such as the government, administration, politics, higher education, the legal system, business and the media (Schneider, 2007). It also functions as an “interethnically neutral link language” (Schneider, 2007: 167), but it is not a marker of identity. The current status of English in India has been described as a “steady state” (Mukherjee, 2007: 158) in which both progressive and conservative forces are at play. The most important progressive forces are the linguistic innovations that distinguish IndE from other varieties, but also certain developments allow us to entertain the possibility that it could become more established in India in the future. Among these are the increase in the number of literary works written in English by Indian authors, and the recent inclusion in the syllabus of a compulsory English subject in primary education. On the other hand,

there are also conservative forces that hinder the spread of English in the country and make it difficult for IndE to become a carrier of Indian identity. First of all, teachers of English follow predominantly a British norm, which prevents the establishment of an Indian standard. Secondly, and most importantly, many Indians consider the innovative features characteristic of IndE as grammatical errors that must be avoided, and instead hold native varieties as the appropriate and correct ones. In terms of Trudgill's (2011) typology of contact situations, IndE is an example of the first type: those characterized by short-term adult L2 acquisition and use.

The most detailed and comprehensive description of relativization strategies in IndE to date is Suárez-Gómez (2014), who focuses on the choice of relativizer in IndE, HKE and SgE restrictive adnominal relative clauses in the spoken component of the *International Corpus of English* (ICE). She found that:⁹

- *who* is favoured with human antecedents in SU relatives, and there are also some cases of *zero* in these contexts;
- *which* is the preferred option with nonhuman referents in SU relatives;
- we find competition between *zero* and *whom* in nonsubject position with human antecedents;
- there is variation between *that*, *zero*, and *which* in nonsubject position with nonhuman referents;
- and in PCOMP relatives, there are very similar frequencies of pied-piping and preposition stranding constructions, and only one instance of preposition deletion.

This distribution shows that spoken IndE is more similar to spoken BrE in the 1950s as described in Quirk (1957), with a preference for *wh*-pronouns over *that* and *zero* and a high frequency of pied-piping constructions in PCOMP relatives. IndE seems to have been unaffected by more recent developments in spoken BrE, which in the 1990s showed a higher frequency of *that* than in the 1950s, and a decrease in the use of *wh*-pronouns (Tottie, 1997). On the other hand, the IndE preference for *wh*-pronouns may also be a reflection of the influence of the substrate languages, particularly Hindi, since relative clauses in this language may be formed by means of relative pronouns, resulting in structures that are very similar to English *wh*-relatives (Suárez-Gómez, 2017).

3.2. Expected distribution of relative clauses in IndE and BrE

Taking into account what was mentioned in sections 2.2 and 3.1, simplification processes are hypothesized to have affected the domain of relative clauses in IndE. IndE is predominantly a L2 variety of English, i.e., a high-contact variety with many adult L2 speakers, and this is expected to be reflected in its syntax, so that IndE contains simpler relative clauses in comparison with BrE, an L1 variety. Simpler relative clauses include, as mentioned in section 2.1, those in which the relativized position is higher in the Accessibility Hierarchy, and those with fewer nodes and levels of hierarchical organization. In the next section (§ 4), this definition of relative clause complexity will

be specified in more detail, together with the data retrieval process and the methodology used in the present study.

4. Data and methodology

The data of the study was extracted from the British (ICE-GB) and Indian (ICE-IND) components of ICE. ICE is a collection of corpora, each component of which contains one million words, 600,000 words of speech and 400,000 of written language, from a variety of English around the world. All the components were compiled following the same annotation template and design, so they are highly comparable. For the purposes of the present paper, 40 texts from each of the two ICE components mentioned above were selected (approximately 82,000 words from ICE-GB and 88,000 from ICE-IND): 10 texts were spoken informal (from the S1A category comprising private conversations), 10 texts were spoken formal (from the S2A category, unscripted speech from broadcast events), 10 were written informal (from the W1B category, social letters), and 10 were written formal (from the W2A category, academic writing). The data in the selected texts was produced in the 1990s: the ICE-GB texts range from the year 1990 to 1993, and those in ICE-IND from 1990 to 1998. Therefore, the differences found between the varieties are not expected to derive from diachronic effects.

All the instances of adnominal relative clauses in the selected texts introduced by a *wh*-pronoun, *that*, or *zero* were retrieved employing different methods:

- Relatives introduced by *wh*-pronouns and *that* in the texts taken from ICE-IND were retrieved using the concordance programme *WordSmith Tools 6*.
- *Zero* relatives in ICE-IND were manually extracted from the texts.
- Relative clauses in ICE-GB were retrieved using *ICECUP 3.1*.

Only instances found in valid data were used, i.e., those retrieved from extra-corpus material (marked <X></X>) were excluded.

The data was then analysed by means of a series of ‘hierarchical configural frequency analyses’ (HCFA), using Gries’ (2004) *HCFA 3.2* script for *R* (R Core Development Team, 2015). HCFA is an extension of the chi-square test that allows the simultaneous analysis of more than 2 variables and that approaches the data in a more exploratory fashion (Hilpert, 2013: 56). This test compares the observed frequencies of the different configurations in a table against the frequency expected by chance: those that are found significantly more often than expected are called *types*, and those that have a significantly lower frequency than expected by chance are known as *antitypes*. HCFA provides the global significance value of the table plus the configurations that are responsible for it. For example, consider Table 1, an extract of the results provided by the HCFA test for one of the analyses conducted in the present study, which is discussed in more depth below.

Variety	Restrictiv.	Text type	Relativized position	Relativizer	Freq	Exp	Cont. chisq	Obs-exp	Dec	Q
BrE	Restrictive	Sp. inf.	SU	wh-pronoun	32	338.817	0.1045	<	ns	0.002
IndE	Restrictive	Sp. inf.	SU	wh-pronoun	22	246.799	0.291	<	ns	0.002
BrE	Restrictive	Sp. inf.	SU	that	29	10.875	302.083	>	***	0.017
IndE	Restrictive	Sp. inf.	SU	that	0	79.215	79.215	<	ms	0.007
BrE	Restrictive	Sp. inf.	SU	zero	0	84.584	84.584	<	*	0.008
IndE	Restrictive	Sp. inf.	SU	zero	0	61.612	61.612	<	ns	0.006

Table 1: Sample of the results of a HCFA test analysing the distribution of relative clauses as a function of variety, restrictiveness, text type, relativized position, and relativizer

The first five columns in the table indicate the levels of each of the variables that are considered in the analysis: BrE and IndE for variety; restrictive and nonrestrictive for restrictiveness; spoken informal, spoken formal, written informal, and written formal for text type; SU, DO, PCOMP, and GEN for relativized position; and *wh*-pronoun, *that*, and *zero* for relativizer. The columns ‘Freq’ and ‘Exp’ provide the observed and expected frequencies respectively of each configuration in the table, and the ‘Cont. chisq’ column gives their chi-square values. ‘Obs-exp’ reflects the relation between the observed and expected frequencies: ‘<’ means less observed frequency than expected, and ‘>’ more than expected. The following column, ‘Dec’, states the significance level of each configuration (‘ns’ = not significant, ‘ms’ = marginally significant, ‘*’ = significant at the 0.05 level, ‘**’ = significant at the 0.01 level, ‘***’ = significant at the 0.001 level). Finally, ‘Q’ stands for coefficient of pronouncedness, a measure of the size of the effect of each configuration (the higher, the stronger). As mentioned above, HCFA also provides global chi-square and significance values for the whole table, which in this case are $\chi^2 = 2010.34$ (d.f. = 181) and $p < 0.001$, i.e., statistically significant.

In the extract of the results provided in Table 1, there are two significant configurations, and one that is marginally significant: restrictive *that* relatives with a SU gap in spoken informal texts are a type in BrE, i.e., they are significantly more frequent than expected by chance; restrictive *zero* relatives with a SU gap in spoken informal texts are an antitype in BrE, i.e., they are less frequent than expected by chance; and restrictive *that* relatives with a SU gap in spoken informal texts are a (marginally significant) antitype in IndE.

4.1. Variables included in the analysis

The following variables (and their levels) were explored by means of HCFA tests:

1. Variety: BrE vs. IndE.
2. Text type: spoken informal vs. spoken formal vs. written informal vs. written formal.
3. Restrictiveness: restrictive vs. nonrestrictive.
4. Relativizer: *wh*-pronoun vs. *that* vs. *zero*.

5. Relativized position: SU (in some cases divided into SU-INT and SU-TR; see below) vs. DO vs. PCOMP vs. GEN.
6. Preposition placement in PCOMP relatives: pied-piping vs. preposition stranding vs. preposition deletion.
7. Relative clause complexity: simple vs. complex.

Variables 1-4 and 6 are simple operationalizations of most of the dimensions of variation in relative clause formation identified in previous research and require no further explanation. As regards variable 5, relativized position, this is an operationalization of the Accessibility Hierarchy and its subsequent refinements. SU position is divided into SU-INT and SU-TR in those tests dealing with relative clause complexity, following Diessel and Tomasello's (2005) findings reviewed in section 2.1. Additionally, the IO position is excluded here because it is not distinguishable from PCOMP in terms of complexity according to Hawkins (1999: 253-254), and because English tends to assimilate IO to PCOMP in relative clause formation¹⁰ (Keenan and Comrie, 1977: 72). Finally, no instances of OC relatives were found in the corpus, so this position is not included in the present study. Examples of relative clauses with each of the relativized positions can be found in (5)-(8).

- (5) Just one and a half months ago <,>uhm<,> I had my aunty with us <,>uhm<,> my aunty [who is above seventy-five] (*SU-INT*) (ICE-IND:S1A-004#162:1:B)
- (6) The line [which attracted me] (*SU-TR*) is this <,> smile is our instrument for winning <,> soul (ICE-IND:S1A-001#157:1:B)
- (7) [...] one of the things [that I felt] (*DO*) when I was studying dance <,> was I very much enjoyed the work [that I was involved in] (*PCOMP*) (ICE-GB:S1A-001 #31:1:B)
- (8) Losing a husband <,> losing a father <,> a loving friend [whose smile could charm a heart of stone] (*GEN*) (ICE-IND:S2A-006#22:1:A)

Variable 7 refers to the complexity of relative clauses and has two levels: simple vs. complex. Complex relatives include those with coordination, as in example (9), and/or further embedding, as in (10):

- (9) His study of Vico, [who denied the knowability of Nature and asserted that of History] (ICE-GB: W2A-003#30:1)
- (10)[...] the opportunity [that has arisen through the group [that we're working with now]] (ICE-GB:S1A-001#38:1:B)

As seen in examples (1), (2), and (3) in section 2, coordinated and embedded clauses are longer than simple ones, with more phrasal nodes and more levels of hierarchical organization: a hypothetical simple clause can already have a 3-level hierarchy and 9 phrasal nodes; coordination may generate a structure with a 4-level hierarchy and 16 nodes; and a sentence with an embedded clause may contain, at least, a 5-level hierarchy and 14 nodes. Therefore, the number of phrasal nodes that has to be parsed increases if the relative clause is coordinated or contains extra dependent clauses. Longer and more

hierarchically embedded relativization domains make the process of gap identification more difficult and, as a consequence, relative clauses become harder to process.¹¹

5. Results

A total of 637 instances of relative clauses in BrE and 464 in IndE were identified in the corpus during the retrieval process and selected for further analysis.¹² Two sets of results are provided in this section based on this data. The first set deals with issues related to relativizer choice and preposition placement in PCOMP relatives. The second set focuses on complexity effects.

5.1. Relativizer choice and preposition placement

Table 2 shows the distribution of relativizers in BrE as a function of restrictiveness, text type, and relativized position.

		RESTRICTIVE			NONRESTRICTIVE		
		<i>wh-pro</i>	<i>that</i>	<i>zero</i>	<i>wh-pro</i>	<i>that</i>	<i>zero</i>
SP. INF.	<i>SU</i>	32	29 T	0 A	16	0	0
	<i>DO</i>	0 A	38 T	8	2	0	0
	<i>PCOMP</i>	3	14 T	15 T	7	0	0
	<i>GEN</i>	0	0	0	0	0	0
SP. FOR.	<i>SU</i>	9 A	8	0 A	69 T	0	0
	<i>DO</i>	1 A	9	4	6	0	0
	<i>PCOMP</i>	1 A	1	1	1	0	0
	<i>GEN</i>	0	1	0	0	0	0
WR. INF.	<i>SU</i>	31	8	1	15	0	0
	<i>DO</i>	2 A	8	42 T	6	0	0
	<i>PCOMP</i>	0 A	3	23 T	3	0	0
	<i>GEN</i>	1	0	0	0	0	0
WR. FOR.	<i>SU</i>	81	34	0 A	26	0 A	0
	<i>DO</i>	12	7	11	4	0	0
	<i>PCOMP</i>	33 T	1	0	4	0	0
	<i>GEN</i>	4	0	0	2	0	0
TOTAL		476			161		

Table 2: Relativizer choice in BrE

Table 2 can be interpreted as follows:

- the number in each cell is the raw frequency of examples in the corpus of a specific configuration (e.g. *wh*-pronouns in *SU* position in spoken informal texts in restrictive relatives);

- the numbers in boldface represent configurations that are statistically significant: the ones followed by ‘T’ are types, and those with an ‘A’ are antitypes (see section 4).

Table 2 does not provide percentages because of the complexity of the data: there are four variables that are hierarchically organized, and three of them have more than two levels. With this kind of data, it is very difficult to decide out of which total the percentages should be calculated. For instance, taking as an example the number of *wh*-pronouns in SU position in spoken informal texts in restrictive relatives, i.e., 32, we could calculate the following percentages, among others, depending on the focus of the study:

- 41.56% out of the total instances of SU relatives in spoken informal texts (77),
- 15.24% out of the total instances of *wh*-pronouns in restrictive relatives (210),
- 8.62% out of the total instances of *wh*-pronouns in both restrictive and nonrestrictive clauses (371),
- 6.72% out of the total instances of restrictive relatives (476),
- 19.51% out of the total instances of relative clauses in spoken informal texts (164).

Since what is of interest here is the general distribution of relativizers in BrE as a function of each and all of the other variables in Table 2, the data can be better summarized in graphical form. Figure 1 is a visual representation of Table 2.

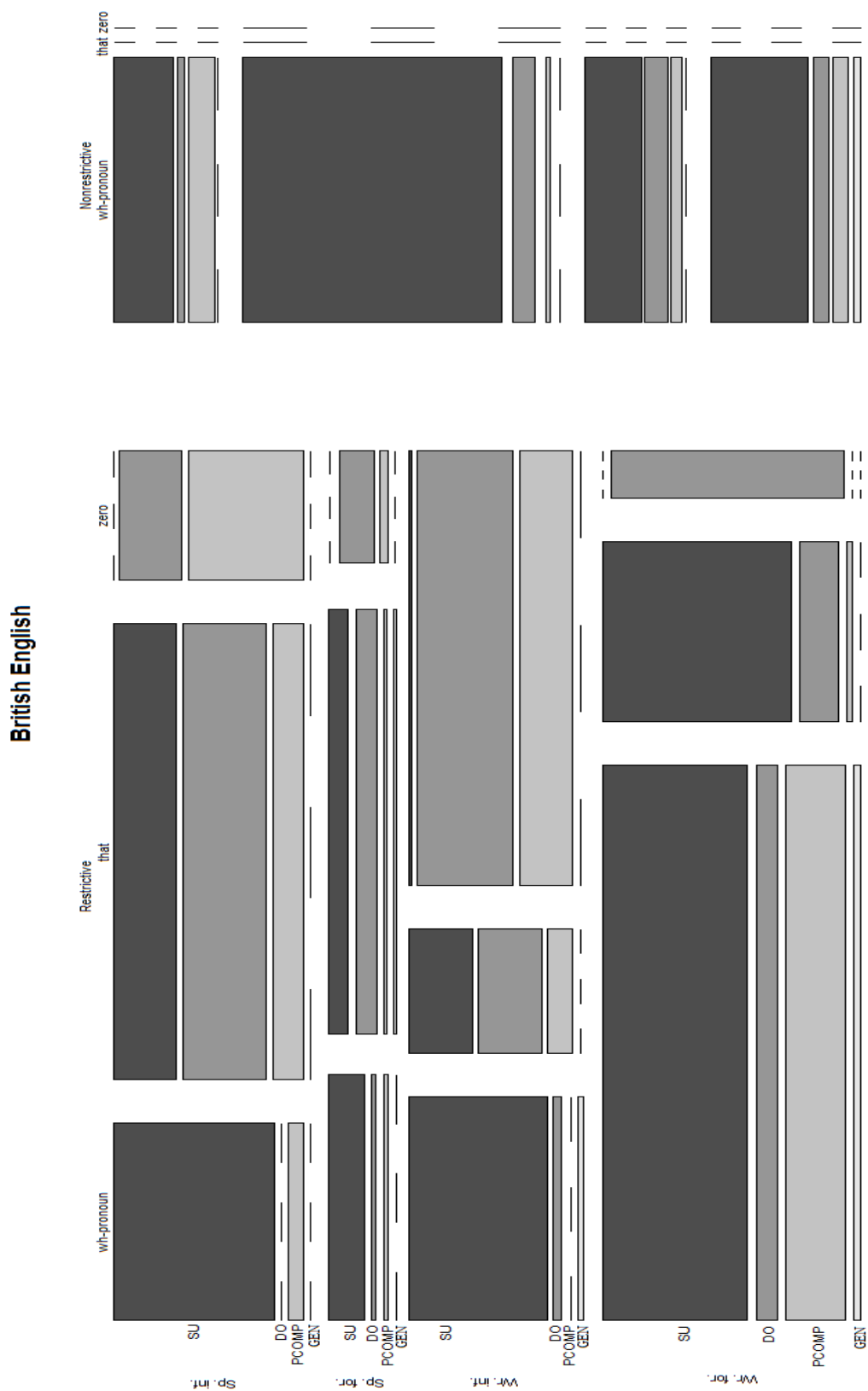


Figure 1: Relativity choice in BrE

Figure 1 plots the proportion of cases of each configuration in relation to the rest, on the basis of the four variables selected: restrictiveness, text type, relativizer, and relativized position. The first division in Figure 1 is that between restrictive and nonrestrictive relative clauses (vertical axis), so that each of the two largest boxes represents the proportion of restrictive and nonrestrictive relatives. Within each of them, we first have a division into text types, and then into relativized positions (horizontal axis): there are four rows that represent the four text types distinguished here and, within each of them, there are four smaller rows, one per relativized position. On the vertical axis there is another division, that between *wh*-pronouns, *that*, and *zero* relativizers. The size of each of the partitions in the plot reflects the proportion of cases of a specific configuration in comparison with the rest. For instance, the black box on the top left corner of Figure 1 represents the proportion of restrictive relative clauses occurring in spoken informal texts with a *wh*-pronoun as a relativizer and with a SU relativized position.

The distribution presented in Table 2 and Figure 1 is statistically significant ($\chi^2 = 2010.34$, d.f. = 181, $p < 0.001$). The picture that emerges from this distribution is a complex one. In restrictive clauses, *that* is favoured in BrE in spoken (both informal and formal) texts, since it is the most frequent option: it is in fact a type in SU, DO, and PCOMP positions in spoken informal texts. *That* occurs more often in SU and DO positions, although it is not an infrequent choice in PCOMP relatives (especially in spoken informal texts). *Zero* is the most frequent relativizer in restrictive relatives in written informal texts. It is favoured in DO and PCOMP positions and strongly disfavoured, or even forbidden, in SU position.¹³ It is a type in PCOMP relatives in spoken informal texts and in DO/PCOMP in written informal ones, and an antitype in SU in spoken informal/formal and written formal texts. *Wh*-pronouns are favoured in restrictive relatives in written formal texts (type in PCOMP position in written formal texts), and disfavoured in the other three text types (antitype in DO in spoken informal texts, in SU/DO/PCOMP positions in spoken formal ones, and in DO/PCOMP relatives in written informal texts). They occur more commonly overall in SU position, although they are also very frequent in PCOMP position in written formal texts, where they are a type. *Wh*-pronouns are the most common option in GEN relatives, which are very infrequent overall. There is, however, one case with *that* and a stranded preposition *of*:

(11)[...] he is carrying this famous letter [that the whole world is waiting to see the contents of] (ICE-GB:S2A-008 #134:3:A)

In nonrestrictive relative clauses, the only available option in BrE are *wh*-pronouns, since there are no instances of *that* or *zero* in the data. They occur more commonly in SU position and in spoken formal texts, where they are a type.

As regards IndE, Table 3 and Figure 2 show the distribution of relativizers in this variety ($\chi^2 = 2010.34$, d.f. = 181, $p < 0.001$):¹⁴

		RESTRICTIVE			NONRESTRICTIVE		
		<i>wh-pro</i>	<i>that</i>	<i>zero</i>	<i>wh-pro</i>	<i>that</i>	<i>zero</i>
SP. INF.	<i>SU</i>	22	0 A	0	9	0	0
	<i>DO</i>	5	4	7	2	0	0
	<i>PCOMP</i>	3	0	7	0	0	0
	<i>GEN</i>	0	0	0	0	0	0
SP. FOR.	<i>SU</i>	21	12	1	55 T	1	0
	<i>DO</i>	3	12 T	8	7	0	0
	<i>PCOMP</i>	3	3	6	1	0	0
	<i>GEN</i>	1	0	0	0	0	0
WR. INF.	<i>SU</i>	17	4	0	31 T	2	0
	<i>DO</i>	3	6	26 T	6	0	0
	<i>PCOMP</i>	4	1	4	6	0	0
	<i>GEN</i>	0	0	0	0	0	0
WR. FOR.	<i>SU</i>	63	12	0 A	31	1	0
	<i>DO</i>	7	5	6	1	1	0
	<i>PCOMP</i>	21	0	5	5	0	0
	<i>GEN</i>	2	0	0	1	0	0
TOTAL		304			160		

Table 3: Relativizer choice in IndE



Figure 2: Relativizer choice in IndE

In IndE restrictive relative clauses, *wh*-pronouns are preferred in spoken informal and written formal texts. They are more frequent in SU position, although no statistical types or antitypes were found by the HCFA test in this case. There is competition between *wh*-pronouns and *that* in spoken formal texts, with a similar number of cases of both relativizers. *That* is disfavoured in spoken informal texts, with only 4 cases attested in the corpus (it is an antitype in SU position in this text type), and it occurs more frequently in DO position overall, becoming a type in spoken formal texts. There is also competition in written informal texts, in this case between *wh*-pronouns and *zero*, though the latter is more frequent. *Zero* is preferred in DO and PCOMP positions (it is a type in DO position in written informal texts), and it is very infrequent in SU relatives¹⁵ (it is an antitype in SU position in written formal texts).

In IndE nonrestrictive relative clauses, *wh*-pronouns are the default choice. They occur more commonly in SU position and in spoken formal and written formal texts, where they are types.¹⁶

With respect to preposition placement in PCOMP relatives, Table 4 and Figure 3 show the distribution of the three different strategies, pied-piping, stranding, and deletion, as a function of restrictiveness and text type in both BrE and IndE ($\chi^2 = 197.94$, d.f. = 40, $p < 0.001$):

		BRITISH ENGLISH		INDIAN ENGLISH	
		<i>Restrictive</i>	<i>Nonrestrictive</i>	<i>Restrictive</i>	<i>Nonrestrictive</i>
SP. INF.	<i>Pied-piping</i>	3 A	2	2	0
	<i>Stranding</i>	14	5	0	0
	<i>Deletion</i>	15 T	0	5	0
SP. FOR.	<i>Pied-piping</i>	1	0	2	1
	<i>Stranding</i>	2	1	3	0
	<i>Deletion</i>	0	0	7 T	0
WR. INF.	<i>Pied-piping</i>	0 A	2	4	6 T
	<i>Stranding</i>	14	1	4	0
	<i>Deletion</i>	12	0	1	0
WR. FOR.	<i>Pied-piping</i>	33 T	4	21 T	5
	<i>Stranding</i>	1 A	0	0	0
	<i>Deletion</i>	0 A	0	3	0
TOTAL		110		64	

Table 4: Preposition placement in BrE and IndE

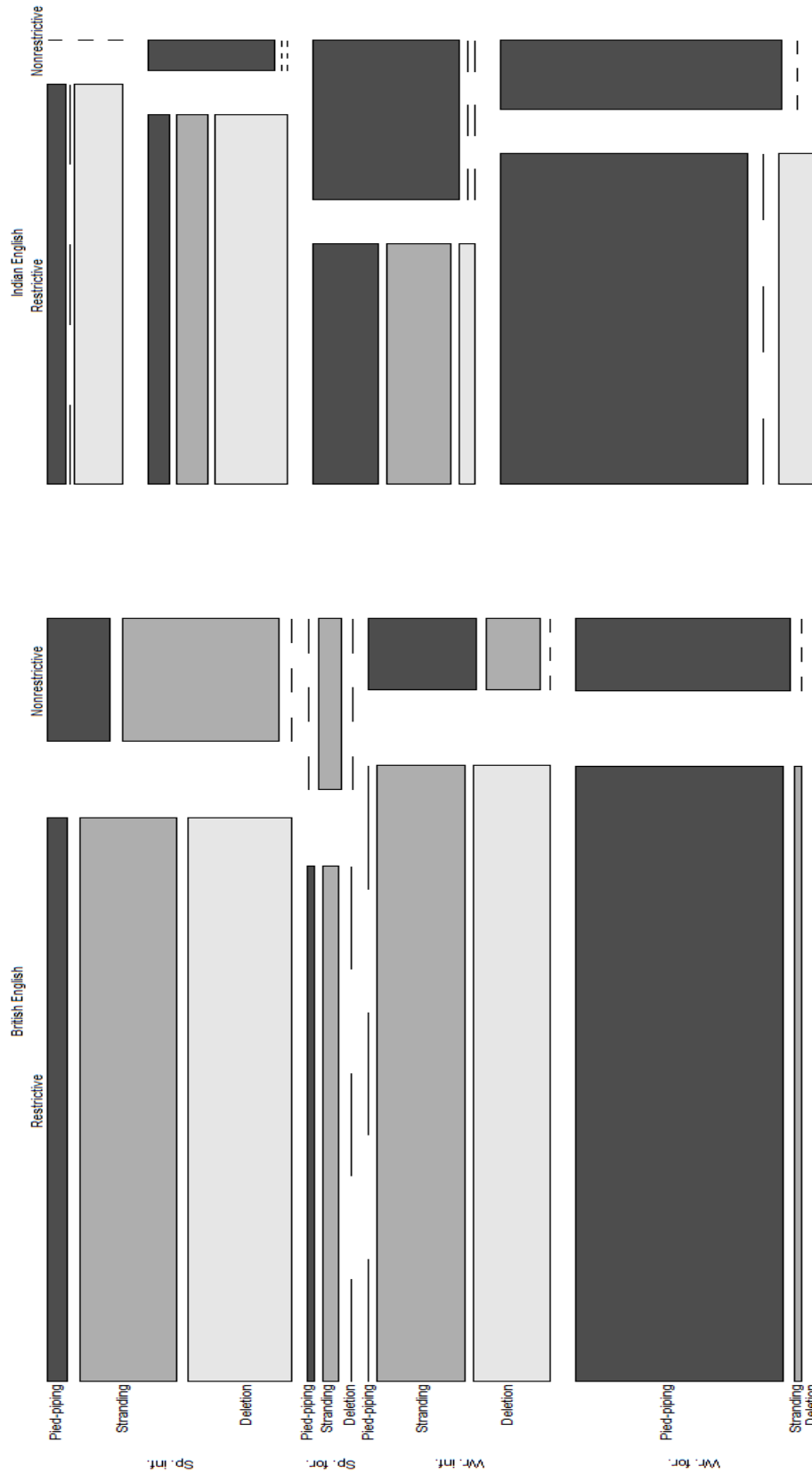


Figure 3: Preposition placement in BrE and IndE

In BrE, pied-piping is favoured in restrictive relatives in written formal texts, where it is a type, and disfavoured (antitype) in spoken informal and written informal texts. Preposition stranding is more common in restrictive clauses in spoken informal and written informal texts, although in this case it is not a statistically significant type, and it is infrequent in written formal ones, in which it becomes an antitype. Similarly, deletion is preferred in restrictive relative clauses in spoken informal (where it is a type) and written informal texts, and disfavoured in written formal ones (where it is an antitype). There are not many cases of nonrestrictive PCOMP relatives, so statistical significant results were not found, but a preference for pied-piping can be observed in written (both formal and informal) texts, while stranding is more common in spoken (again, both formal and informal) ones.

In IndE, pied-piping is more common in restrictive clauses in written formal texts, in which it is a type, and it is in competition with stranding in written informal ones. In spoken (both informal and formal) texts, on the other hand it is disfavoured, although it is not a statistically significant antitype. Stranding is only frequent in IndE restrictive relatives in written informal texts, where, as just mentioned, it is in competition with pied-piping. Deletion is more common in spoken texts, being a type in formal ones, and disfavoured in written (both informal and formal) texts. With respect to nonrestrictive relative clauses, pied-piping is the only variant attested in IndE in the data, and it is especially frequent in written informal texts, where it is a type.

5.2. Complexity effects

Tables 5 and 6 show the individual effects of relativized position ($\chi^2 = 5.95$, d.f. = 4, $p > 0.05$) and relative clause complexity ($\chi^2 = 6.34$, d.f. = 1, $p < 0.05$) respectively.

	BRITISH ENGLISH	INDIAN ENGLISH
SU	233 (48.95%)	152 (50%)
<i>SU-INT</i>	97 (20.38%)	82 (26.97%)
<i>SU-TR</i>	136 (28.57%)	70 (23.03%)
DO	142 (29.83%)	92 (30.26%)
PCOMP	95 (19.96%)	57 (18.75%)
GEN	6 (1.26%)	3 (0.99%)
TOTAL	476 (100%)	304 (100%)

Table 5: Frequency of relative clauses per relativized position in BrE and IndE¹⁷

	BRITISH ENGLISH	INDIAN ENGLISH
SIMPLE	332 (69.75%)	237 (77.96%)
COMPLEX	144 (30.25%)	67 (22.04%)
TOTAL	476 (100%)	304 (100%)

Table 6: Frequency of simple and complex relative clauses in BrE and IndE

Even though the global distribution of simple and complex relative clauses in BrE and IndE is statistically significant, we do not find any significant types or antitypes with respect to the variables relativized position and relative clause complexity, which means that they do not have individual effects on the distribution in the varieties at hand. We find that, in both varieties, there are more instances of relative clauses in SU position (SU-INT + SU-TR), then in DO, PCOMP, and, finally, in GEN position. In IndE, there are also more cases of SU-INT than SU-TR relatives, whereas BrE shows the opposite distribution: SU-TR > SU-INT. With respect to relative clause complexity, simple relatives are much more frequent than complex ones.

Statistically significant results are found when we investigate the conjoined effect of both variables. Table 7 and Figure 4 show the interaction of relativized position and relative clause complexity in BrE and IndE ($\chi^2 = 36.05$, d.f. = 13, $p < 0.001$).

	BRITISH ENGLISH		INDIAN ENGLISH	
	<i>Simple</i>	<i>Complex</i>	<i>Simple</i>	<i>Complex</i>
SU	149 (31.31%)	84 (17.64%)	112 (36.84%)	40 (13.16%)
<i>SU-INT</i>	65 (13.66%)	32 (6.72%)	61 (20.06%)	21 (6.91%)
<i>SU-TR</i>	84 (17.65%)	52 T (10.92%)	51 (16.78%)	19 (6.25%)
DO	107 (22.48%)	35 (7.35%)	69 (22.70%)	23 (7.56%)
PCOMP	74 (15.55%)	21 (4.41%)	53 (17.43%)	4 A (1.32%)
GEN	2 (0.42%)	4 (0.84%)	3 (0.99%)	0 (0%)
TOTAL	476 (100%)		304 (100%)	

Table 7: Conjoined effect of relativized position and relative clause complexity in BrE and IndE

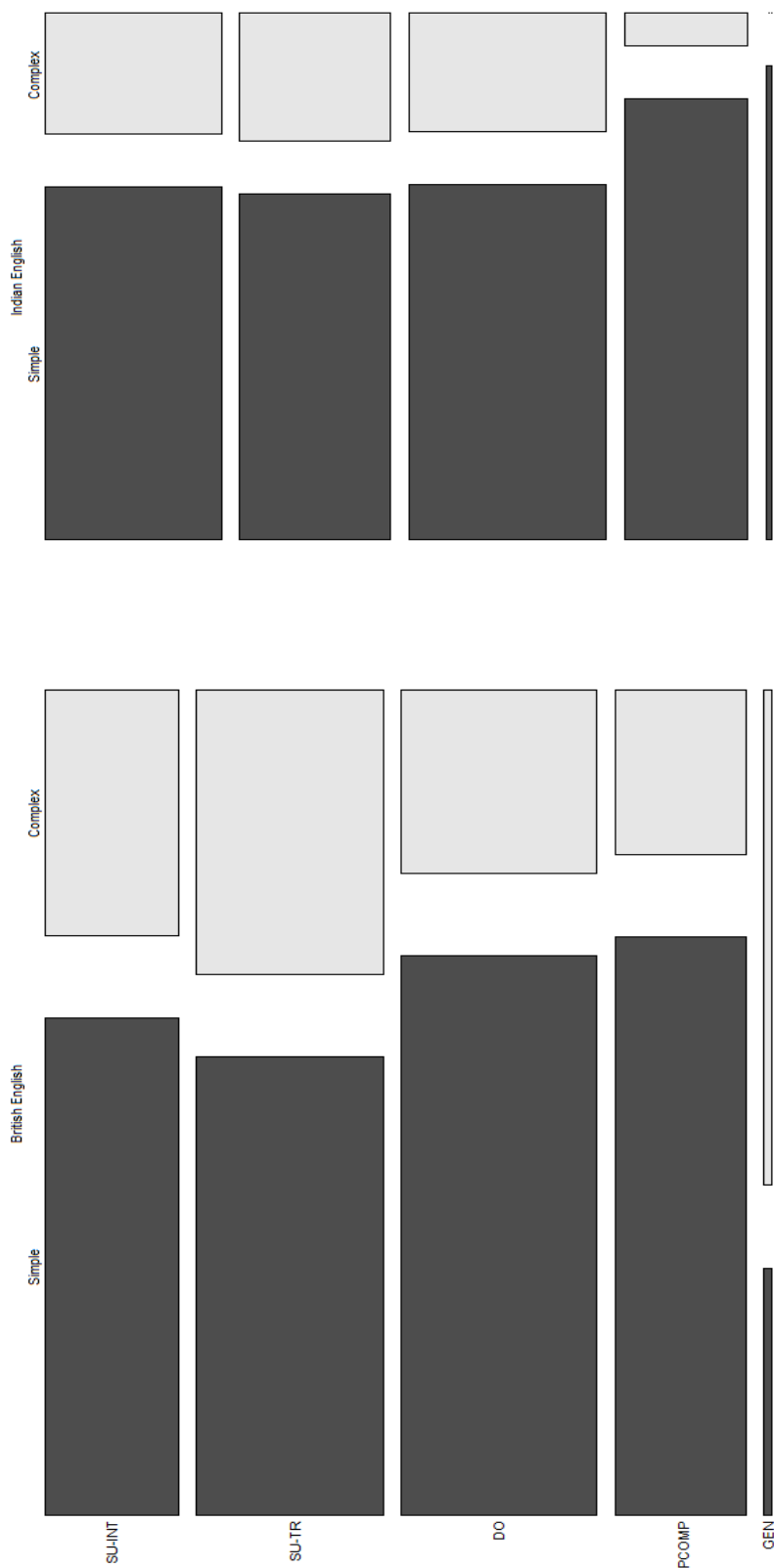


Figure 4: Conjoined effect of relativized position and relative clause complexity in BrE and IndE

There are two interesting statistically significant configurations: complex relatives are a type in SU-TR position in BrE, and an antitype in PCOMP position in IndE. No significant results were found in GEN position, probably due to the low number of instances, but the distribution is suggestive: there are no cases of GEN complex relatives in IndE, while in BrE they are more frequent than simple ones. Both varieties comply with the Accessibility Hierarchy, with fewer instances of simple and complex relatives in the lower positions. However, we do find differences between them: complex relatives are more strongly disfavoured in PCOMP and GEN positions in IndE, i.e., in those in which the process of gap identification is more difficult, and they are more strongly favoured in SU-TR in BrE, i.e., in the SU position in which forming relatives is more complicated. This distribution points to the conclusion that relative clauses are simpler in IndE than in BrE, an issue that is further discussed in section 6.2.

6. Discussion

6.1. Relativizer choice and preposition placement

As seen in section 5.1, BrE shows a preference for covert relativizers (*that* and *zero*) in restrictive relative clauses: *that* is favoured in speech (both informal and formal) and SU/DO positions, and *zero* in written informal texts and DO/PCOMP positions. *Wh*-pronouns are favoured in written formal texts and in SU/GEN positions in restrictive relatives, and they are the only relativizers used in nonrestrictive ones. This distribution agrees with previous descriptions of relativizer choice in English, with *wh*-pronouns being preferred in SU position in written formal contexts and in nonrestrictive relatives, and *that* and *zero* in speech and informal texts. In IndE, on the other hand, overt relativizers, i.e., *wh*-pronouns, are the preferred choice: in restrictive relatives, they are the most frequent option in spoken informal and written formal texts, and they are in competition with *that* and *zero* in spoken formal and written informal ones respectively. They are again favoured in SU and GEN positions, and they are not infrequent in PCOMP relatives in written formal contexts. *That* is only frequent in spoken formal texts and DO position, while *zero* is favoured only in written informal texts and DO/PCOMP positions. As in BrE, *wh*-pronouns are the default option in nonrestrictive relative clauses, with only a few cases of marginally non-restrictive *that* relatives. Overall, these results agree with Suárez-Gómez (2014), with a preference for *wh*-pronouns in SU position and competition between *that*, *zero*, and *wh*-pronouns in non-SU positions, with *wh*-pronouns being the most common relativizers overall.¹⁸ These findings, as argued by Suárez-Gómez (2014: 259), characterize IndE as being more formal and similar to educated BrE in the 1950s. Additionally, the IndE preference for *wh*-pronouns over other relativizers may be the result of substrate influence (see § 3.1).

There are two findings in the present study that are unexpected and require further explanation: the high frequencies of *zero* relatives in written informal texts (in both IndE and BrE) and *that* relatives in spoken formal ones (in IndE). With respect to the former, *zero* relativizers, as has been mentioned before (§ 3), do tend to occur in informal

registers, but why should they be mostly restricted to written language? The *zero* relativizer is the variant that provides the least information: there is no relative marker at the beginning of the embedded clause, contrary to *wh*-pronoun and *that* relatives, which do have an explicit marker, and there is no information about the function of the gap in the clause, which is provided in *wh*-pronoun relatives in most cases. As a consequence, identifying the relative clause and the gap within it is harder for the addressee in *zero* relatives. In written language, however, the temporal constraints on spoken communication can be ignored, since the reader has access to the previous discourse and does not have the same pressure of keeping information in short-term memory (Mass, 2009: 166). Therefore, the lack of an explicit marker in *zero* relatives does not cause as many problems for the addressee in written language as it does in speech. As regards the frequency of *that* relatives in spoken formal language in IndE, this is indeed a very unexpected finding, taking into account what has been found in previous studies (§ 3.1), and is still in need of further research.

With respect to preposition placement in PCOMP relatives, BrE favours pied-piping in written formal texts in restrictive clauses, and there is competition between stranding and deletion in informal registers (both spoken and written). In nonrestrictive relatives, pied-piping is preferred in writing, and stranding in speech. These findings are only in partial agreement with previous research: pied-piping is not disfavoured in nonrestrictive relatives, although it is indeed more frequent in written formal contexts while stranding and deletion dominate in informal registers. In IndE, there is a higher frequency of pied-piping overall than in BrE: it is favoured in writing in restrictive relatives, and it is the only structure in nonrestrictive ones. Deletion is the preferred option in speech, and stranding is only frequent in written informal texts, where it is in competition with pied-piping. These results are again not completely aligned with those of previous studies, since deletion is not infrequent in speech (although it is rare overall) and stranding is only common in written informal texts. Pied-piping is preferred only in written formal texts and nonrestrictive relative clauses.

On the whole, there is more variation in restrictive relatives than in nonrestrictive ones. With respect to the choice of relativizer, there is more competition between the different alternatives in restrictive clauses, while in nonrestrictive ones *wh*-pronouns seem to be the default option. As regards the position of the preposition in PCOMP relative clauses, pied-piping is the only structural option available in nonrestrictive relatives in IndE, and a very common one in BrE. The situation again is more complex in restrictive relatives, with more competition between pied-piping, stranding, and deletion in both varieties.

6.2. Complexity effects

Both varieties follow the Accessibility Hierarchy proposed by Keenan and Comrie (1977): we find fewer instances of relative clauses as we go down the hierarchy due to the increased difficulty of relativizing a NP in the lower positions. With respect to the complexity of relative clauses, there is also a preference for simple relatives in both BrE and IndE, rather than for more complex ones with coordination or further embedding.

The most interesting findings, however, emerge from the interaction between the two variables relativized position and relative clause complexity: complex relatives decrease in frequency in the lower positions in the Accessibility Hierarchy, and this tendency is stronger in IndE than in BrE, with almost no cases of complex relatives in PCOMP and GEN positions in the former (4 in PCOMP and 0 in GEN). Additionally, in BrE there are many cases of complex relatives in SU-TR position, i.e., the SU position with which children had more problems in Diessel and Tomasello's (2005) study. It seems that BrE speakers opt for more complicated SU relatives (complex SU-TR) more often than expected by chance, a tendency that is not found in IndE.

The differences in relative clause formation between BrE and IndE can be characterized as differences in complexity. IndE speakers tend to produce simpler relative clauses than BrE speakers, with simpler domains for the processing of head NP-gap dependencies: they disfavour complex relatives in PCOMP and GEN positions, i.e., relative clauses with coordination and further embedding, constructions which, as argued in section 4, increase the number of nodes in the way from the head NP to the gap. PCOMP and GEN are the most difficult to relativize, and, therefore, IndE speakers tend to produce syntactically simpler and shorter relatives in these positions. The added complexity of relativizing a PCOMP or GEN NP and producing a complex clause seems to be too costly for these speakers, who, as mentioned in section 3.1, are mostly L2 or L3 speakers of English with non-native proficiency. The findings of the present study thus reinforce what has been found before in the literature: language contact that is characterized by short-term adult acquisition has a simplifying influence on languages/dialects.

7. Conclusions

This study has provided interesting results with respect to, on the one hand, relativizer choice and preposition placement, and, on the other, relative clause complexity in BrE and IndE. As concerns the selection of the relativizer and the position of the preposition in PCOMP relatives, the inclusion of different text types and nonrestrictive relative clauses in the analysis yielded a clearer picture of the variation found in the data. In line with previous research (cf., for instance, Gut and Coronel, 2012), text type turned out to be very important, since some relativizers are more characteristic of certain text types and not others: *zero* relativizers were mostly found in written informal texts in both BrE and IndE, a finding that was explained in terms of register and processing considerations (§ 6.1). Other distributions encountered in the present study, such as the unexpected high frequency of *that* relatives found in spoken formal texts in IndE, are still in need of further research. With regard to restrictiveness, more variation and competition between the different structural options were found in restrictive relatives. Nonrestrictive clauses, on the contrary, seem to be more homogeneous.

The analysis of relativizer choice and preposition placement is also significant for another reason: it can be used to test the reliability of the results of the present study against the background of previous research on relative clause formation in varieties of

English, which has so far been mostly focused on these two issues. The data for the present article has been extracted from a relatively small corpus (fewer than 90,000 words per variety), which casts doubt on the generalizability of the results. However, the distribution of relativizers and structural options regarding preposition placement in PCOMP relatives in the present data is very similar to that found in previous research based on larger corpora (§ 3), which suggests that the results of the present study are indeed worthy of consideration.

The main focus of the article lay, however, on complexity effects in the domain of relative clause formation in BrE and IndE. The results replicate a pattern that has been previously found in the literature: contact in which short-term adult L2 acquisition dominates tends to simplify languages. In this case, IndE, a high-contact L2 variety, shows a preference for simpler relative clauses in comparison with BrE. In order to properly understand this finding, it is crucial to consider the manner in which complexity was measured. The metric used here was a structural one, i.e., it measured the complexity of individual structures. Therefore, we cannot claim that the syntax of IndE is simpler than that of BrE in the domain of relative clauses, but it can be stated that IndE speakers tend to produce relative clauses that are easier to process, and that this is a consequence of their condition of L2/L3speakers of English. The metric used in the present study was also a relative one, i.e., it was rooted on the preferences of language users. Contrary to absolute metrics which, by definition, are independent from considerations of language use, relative metrics can locate the source of the complexity differences between BrE and IndE, which in this case emerge from the cost of processing relative constructions. IndE speakers, being non-native users of English, have more problems than BrE speakers when it comes to producing and comprehending relative clauses that relativize a PCOMP or GEN NP and that contain coordinate clauses and/or further embedding, due to the added processing cost. This, in turn, suggests that at least part of the decrease in complexity that takes place in contact situations is not due to imperfect L2 acquisition, but a result of structural simplifications that originate in language use. Thus, in future research, performance effects should be distinguished from simplifications that stem from the process of adult language learning.

Acknowledgements

Some of the material in this paper was originally presented at the *Change 2015* conference held at the University of Helsinki, 8-10 June 2015; I would like to thank the participants at the conference for helpful discussion. For generous financial support, I am grateful to the European Regional Development Fund and the following institutions: Regional Government of Galicia (Directorate General for Scientific and Technological Promotion, grant GPC2014/004); Spanish Ministry of Economy and Competitiveness (grants FFI2014-52188-P and BES-2015-071233). Finally, I would like to thank two anonymous reviewers for their useful comments and suggestions.

I declare that the work presented in this paper is original, and that it does not copy or reproduce the work of any other person, except as acknowledged in the references.

Notes

1. This does not mean that all the differences found in the present study between BrE and IndE with respect to the domain of relative clauses can be attributed to contact. While contact does in fact exert a strong influence in language variation, it is not its only cause; substrate effects, for instance, are another important motivation for change. I would like to thank an anonymous reviewer for drawing attention to this issue.

2. Syntactic complexity has also been operationalized in terms of the length in words or syllables of a syntactic unit (Wasow, 1997; Arnold et al., 2000; Wasow and Arnold, 2003). However, as demonstrated by Szmezcanyi (2004), this metric is highly correlated with the number of nodes that a syntactic unit contains.

3. As an anonymous reviewer points out, relativizers differ in how complex/simple they are. For example, in the English relativization system, *that* relatives are considered to be simpler than *zero* relatives because the former are more transparent than the latter: in *zero* relatives the relativizer is not explicit. Furthermore, *that* is simpler than *wh*-pronouns, since these must agree with the animacy of the antecedent (*who(m)* with animate and *which* with inanimate antecedents) while *that* is invariable in this respect. Agreement adds redundancy to a structure because it implies that one meaning is expressed by means of two or more forms and, therefore, it increases complexity (cf. Suárez-Gómez, 2017). While these issues are very important to account for the complexity of relative clauses, the present paper focuses on complexity from a syntactic perspective.

4. Indirect objects and prepositional complements are not different with respect to their complexity in the metric proposed by Hawkins (1999).

5. Objects of comparison are excluded from Hawkins' (1999: 253) metric because "the coding of this position is highly variable across languages".

6. Many studies found support for the Accessibility Hierarchy in both L1 and L2 acquisition (cf. Izumi, 2003, and references therein).

7. Relative adverbs (*where*, *when*, and *why*) also occur in English when the gap is an adjunct, but these forms lie out of the scope of the present study.

8. *Zero* does occur in subject position in a restricted set of constructions, although very infrequently (Quirk et al., 1985: 1250; Biber et al., 1999: 619; Huddleston and Pullum et al., 2002: 1055).

9. Adnominal relative clauses are those "which depend on an explicitly mentioned nominal antecedent" (Suárez-Gómez, 2014: 246).

10 For example, in *I met the man to whom you gave the book*, the relative pronoun functions as the complement of the preposition *to*. No cases of actual IO relatives (e.g., *I met the man whom you gave the book*) were found.

11. As an anonymous reviewer observes, the complexity of the relative clause is not the only factor that influences the processing cost of relative structures. Another important issue to take into account is the position of the relative in the main clause (Izumi, 2003, and references therein; Diessel and Tomasello, 2005). Thus, relative clauses embedded in head NPs functioning as subjects of the main clause (as in example (6) above) are harder to process than those embedded in a nonsubject NP (as in (5)), because they hinder the parsing of the sentence by adding extra material between the subject and the main verb. This factor was explored in an initial stage of the investigation but was then abandoned since no significant differences between BrE and IndE were found.

12. The fact that we find fewer cases of relative clauses in IndE than in BrE in a similar number of words may already be an indication of the simplification processes at play in the former variety in this grammatical domain. Relative clauses are hard to process (see § 2.1) and, therefore, it is not surprising that we find fewer instances of this structure in the L2 variety.

13. There is one case of a *zero* relativizer in SU position in BrE. It functions as the subject of an embedded clause, a context in which relativizer omission in SU is allowed in English (cf. Huddleston and Pullum et al., 2002: 1047): *She played her “Minstrel Showboat” the one [ø you said sounded Chinese] for about 100 people [...]* (ICE-GB: W1B-007 #114:3).

14. The global chi-square results are the same for BrE and IndE because variety was another variable in the HCFA test, i.e., the data from both varieties was included in the test in order to compare them.

15. There is also one case of a *zero* relativizer in SU in IndE. It occurs in the same type of context as the example found in BrE (see endnote 13): *The other point [ø I think is <, > important] Korean going to do a lot of interceptions* (ICE-IND:S2A-004#96:1:D).

16. There are five cases of *that* relatives in IndE that are not prototypical examples of restrictive clauses and were classified as nonrestrictive, as in *The whole misunderstanding about Hume's philosophical position is the outcome of his treatment of causation [that is often misunderstood]* (ICE-IND: W2A-001#58:1). However, as suggested by Denison and Hundt (2013: 162), a binary distinction between restrictive and nonrestrictive relatives might not be the best way to classify the data. As the focus of the present study is not on how to categorize relative clauses with respect to the relation between the head NP and the relative clause, these examples are not discussed further.

17. Only restrictive clauses are taken into account in the analysis of complexity effects. The syntactic relation between the head NP and the relative clause is different in restrictive and nonrestrictive relatives (cf. Huddleston and Pullum et al., 2002: 1058), and, therefore, processing the dependency between the head NP and the gap may also be different.

18. Suárez-Gómez's (2014) study focuses on restrictive relative clauses in the private spoken component of ICE-IND.

Primary sources

ICE-GB = *International Corpus of English - the British Component* (1990). Project Coordinated by Prof. Gerald Nelson at the Chinese University of Hong Kong. URL: <<http://www.ice-corpora.net/ice/download.htm>>.

ICE-India = *International Corpus of English - the Indian Component* (2002). Project Coordinated by Prof. S. V. Shastri at Shivaji University and Prof. Dr. Gerhard Leitner at FreieUniversität Berlin. URL: <<http://www.ice-corpora.net/ice/download.htm>>.

References

- Arnold, Jennifer, Thomas Wasow, Anthony Losongco and Ryan Ginstrom (2000): “Heaviness vs. newness: The effects of structural complexity and discourse status on constituent ordering”. *Language*, 76(1): 28-55.
- Biber, Douglas; Stig Johansson, Geoffrey Leech, Susan Conrad and Edward Finegan (1999): *Longman Grammar of Spoken and Written English*. Essex: Longman.

- Dahl, Östen (2004): *The Growth and Maintenance of Linguistic Complexity*. Amsterdam: John Benjamins.
- Dahl, Östen (2009): "Testing the assumptions of complexity invariance: The case of Elfdalian and Swedish". In G. Sampson, D. Gil and P. Trudgill, ed., *Language Complexity as an Evolving Variable*. Oxford: Oxford University Press, 50-63.
- Denison, David and Marianne Hundt (2013): "Defining relatives". *Journal of English Linguistics*, 41(2): 135-167.
- Diessel, Holger (2009): "On the role of frequency and similarity in the acquisition of subject and non-subject relative clauses". In T. Givón and M. Shibatani, ed., *Syntactic Complexity: Diachrony, Acquisition, Neuro-cognition, Evolution*. Amsterdam: John Benjamins, 251-276.
- Diessel, Holger and Michael Tomasello (2005): "A new look at the acquisition of relative clauses". *Language*, 81(1): 1-25.
- Givón, Talmy (2009): *The Genesis of Syntactic Complexity: Diachrony, Ontogeny, Neuro-Cognition, Evolution*. Amsterdam: John Benjamins.
- Gries, Stefan Th. (2004): *HCFA 3.2. A program for R*. Available at: <<http://www.linguistics.ucsb.edu/faculty/stgries/>>.
- Gut, Ulrike and Lilian Coronel (2012): "Relatives worldwide". In M. Hundt and U. Gut, ed., *Mapping Unity and Diversity World-Wide: Corpus-based Studies of New Englishes*. Amsterdam: John Benjamins, 215-241.
- Hawkins, John A. (1994): *A Performance Theory of Order and Constituency*. Cambridge: Cambridge University Press.
- Hawkins, John A. (1999): "Processing complexity and filler-gap dependencies across grammars". *Language*, 75(2): 244-285.
- Hawkins, John A. (2004): *Efficiency and Complexity in Grammars*. Oxford: Oxford University Press.
- Hilpert, Martin (2013): *Constructional Change in English: Developments in Allomorphy, Word Formation, and Syntax*. Cambridge: Cambridge University Press.
- Hofmann, Thomas (2005): "Variable vs. categorical effects: Preposition pied piping and stranding in British English relative clauses". *Journal of English Linguistics*, 33(3): 257-297.
- Huddleston, Rodney, Geoffrey K. Pullum, Laurie Bauer, Betty Birner, Ted Briscoe, Peter Collins, David Denison, David Lee, Anita Mittwoch, Geoffrey Nunberg, Frank Palmer, John Payne, Peter Peterson, Lesley Stirling and Gregory Ward (2002): *The Cambridge Grammar of the English Language*. Cambridge: Cambridge University Press.
- Izumi, Shinichi (2003): "Processing difficulty in comprehension and production of relative Clauses by learners of English as a second language". *Language Learning*, 53(2): 285-323.
- Karlsson, Fred (2009): "Origin and maintenance of clausal embedding complexity". In G. Sampson, D. Gil and P. Trudgill, ed., *Language Complexity as an Evolving Variable*. Oxford: Oxford University Press, 192-202.
- Keenan, Edward and Bernard Comrie (1977): "Noun phrase accessibility and universal grammar". *Linguistic Inquiry*, 8(1): 63-99.
- Keenan, Edward and Bernard Comrie (1979): "Data on the noun phrase accessibility hierarchy". *Language*, 55(2): 333-351.
- Kortmann, Bernd and Benedikt Szmrecsanyi (2009): "World Englishes between simplification and complexification". In L. Siebers and T. Hoffmann, ed., *World Englishes – Problems, Properties and Prospects: Selected Papers from the 13th IAWC Conference*. Amsterdam: John Benjamins, 265-285.

- Kortmann, Bernd and Benedikt Szmrecsanyi (2011): "Parameters of morphosyntactic variation in World Englishes: Prospects and limitations of searching for universals". In P. Siemund, ed., *Linguistic Universals and Language Variation*. Berlin/New York: Mouton de Gruyter, 264-290.
- Kusters, Wouter (2003): *Linguistic Complexity: The Influence of Social Change on Verbal Inflection*. Utrecht: LOT.
- Kusters, Wouter (2008): "Complexity in linguistic theory, language learning and language change". In M. Miestamo, K. Sinnemäki and F. Karlsson, ed., *Language Complexity: Typology, Contact, Change*. Amsterdam: John Benjamins, 3-22.
- Maas, Utz (2009): "Orality versus literacy as a dimension of complexity". In G. Sampson, D. Gil and P. Trudgill, ed., *Language Complexity as an Evolving Variable*. Oxford: Oxford University Press, 164-177.
- Matras, Yaron (2009): *Language Contact*. Cambridge: Cambridge University Press.
- McWhorter, John H. (2001): "The world's simplest grammars are creole grammars". *Linguistic Typology*, 5(2-3): 125-166.
- McWhorter, John H. (2007): *Language Interrupted: Signs of Non-native Acquisition in Standard Language Grammar*. Oxford: Oxford University Press.
- Miestamo, Matti (2008): "Grammatical complexity in a cross-linguistic perspective". In M. Miestamo, K. Sinnemäki and F. Karlsson, ed., *Language Complexity: Typology, Contact, Change*. Amsterdam: John Benjamins, 23-41.
- Miestamo, Matti; Kaius Sinnemäki, and Fred Karlsson (2008): *Language Complexity: Typology, Contact, Change*. Amsterdam: John Benjamins.
- Mukherjee, Joybrato (2007): "Steady states in the evolution of New Englishes: Present-day Indian English as an equilibrium". *Journal of English Linguistics*, 35(2): 157-187.
- Mukherjee, Joybrato and Stefan Gries (2009): "Collostructional nativization in World Englishes: Verb-construction associations in the International Corpus of English". *English World-Wide*, 30(1): 27-51.
- Parkvall, Mikael (2008): "The simplicity of creoles in a cross-linguistic perspective". In M. Miestamo, K. Sinnemäki and F. Karlsson, ed., *Language Complexity: Typology, Contact, Change*. Amsterdam: John Benjamins, 265-285.
- R Core Development Team (2015): *R: A Language and Environment for Statistical Computing*. Vienna: The R Foundation for Statistical Computing. URL: <<http://www.R-project.org>>.
- Quirk, Randolph (1957): "Relative clauses in educated spoken English". *English Studies*, 38(1-6): 97-109.
- Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech and Jan Svartvik (1985): *A Comprehensive Grammar of the English Language*. Essex: Longman.
- Sharma, Devyani (2010): "Indian English". In B. Kortmann and K. Lunkenheimer, eds., *The Mouton World Atlas of Variation in English*. Berlin/New York: Mouton de Gruyter, 523-530.
- Schneider, Edgar (2007): *Postcolonial English: Varieties around the World*. Cambridge: Cambridge University Press.
- Schröter, Verena and Bernd Kortmann (2016): "Pronoun deletion in Hong Kong English and colloquial Singaporean English". *World Englishes*, 35(2): 221-241.
- Sinnemäki, Kaius (2009): "Complexity in core argument marking and population size". In G. Sampson, D. Gil and P. Trudgill, eds., *Language Complexity as an Evolving Variable*. Oxford: Oxford University Press, 126-140.
- Suárez-Gómez, Cristina (2014): "Relative clauses in Asian Englishes". *Journal of English Linguistics*, 42(3): 245-268.

- Suárez-Gómez, Cristina (2017): "Transparency and language contact in the nativization of relative clauses in World English". *English World-Wide*, 38(2): 211-237.
- Szmrecsanyi, Benedikt (2004): "On operationalizing syntactic complexity". In G. Purnelle, C. Fairon and A. Dister, ed., *Le poids des mots. Proceedings of the 7th International Conference on Textual Data Statistical Analysis. Louvain-la-Neuve, March 10-12, 2004*. Louvain-la-Neuve: Presses Universitaires de Louvain, 1032-1039.
- Szmrecsanyi, Benedikt and Bernd Kortmann (2009a): "The morphosyntax of varieties of English worldwide: A quantitative perspective". *Lingua*, 119(11): 1643-1663.
- Szmrecsanyi, Benedikt and Bernd Kortmann (2009b): "Between simplification and complexification: Non-standard varieties of English around the world". In G. Sampson, D. Gil and P. Trudgill, ed., *Language Complexity as an Evolving Variable*. Oxford: Oxford University Press, 64-79.
- Szmrecsanyi, Benedikt and Bernd Kortmann (2012): "Introduction: linguistic complexity – Second language acquisition, indigenization, contact". In B. Kortmann and B. Szmrecsanyi, ed., *Linguistic Complexity: Second Language Acquisition, Indigenization, Contact*. Berlin/Boston: Walter de Gruyter, 6-34.
- Thomason, Sarah and Terrence Kaufmann (1988): *Language Contact, Creolization, and Genetic Linguistics*. Berkeley: University of California Press.
- Tottie, Gunnel (1997): "Relatively speaking: Relative marker usage in the British National Corpus". In T. Nevalainen and L. Kahlas-Tarkka, ed., *To Explain the Present: Studies in the Changing English Language in Honour of Matti Rissanen*. Helsinki: Société Néophilologique, 465-481.
- Trudgill, Peter (2009): "Sociolinguistic typology and complexification". In G. Sampson, D. Gil and P. Trudgill, ed., *Language Complexity as an Evolving Variable*. Oxford: Oxford University Press, 98-109.
- Trudgill, Peter (2011): *Sociolinguistic Typology: Social Determinants of Linguistic Complexity*, Oxford: Oxford University Press.
- Valdés, Guadalupe (2012): "Multilingualism". *Linguistic Society of America*. Available at: <<http://www.linguisticsociety.org/resource/multilingualism>> [accessed 24 November 2017].
- Wasow, Thomas (1997): "Remarks on grammatical weight". *Language Variation and Change*, 9(1): 81-105.
- Wasow, Thomas and Jennifer Arnold (2003): "Post-verbal constituent ordering in English". In G. Rohdenburg and B. Mondorf, ed., *Determinants of Grammatical Variation in English*. Berlin/New York: Mouton de Gruyter, 119-154.

Part II

Exploring trends and methodologies in the
study of anglicized lexis

English lexical items in Egyptian Arabic. Code-switching or borrowing?

Małgorzata Kniaź
Jagiellonian University
malgorzata.kniaz@gmail.com

ABSTRACT

The study aims to answer the question of whether lone English items that occur in Arabic-English code-switching are borrowings or code-switches. This is based on empirical data collected at the American University in Cairo. The data were analyzed within the framework of the Matrix Language Frame model. 3443 bilingual projections of complementizer (CP) were investigated. They were divided into two types: (1) CPs with Arabic as the Matrix Language (ML) and (2) CPs with English as the ML. The analysis shows a clear discrepancy between categories of items used in the two types. In Arabic CPs, the most frequently switched category concerns English nouns related to the field of study and academic life as well as Standard Arabic in monolingual discourse. The interviews conducted with the participants in the study revealed that they were mostly used due to the lack of Arabic equivalents at the speakers' disposal. In English CPs, mainly Arabic conjunctions and discourse markers appear. We claim that this categorial and functional variation between Arabic and English results from the linguistic situation in the Arab world. For bilingual speakers in Arabic diglossic communities, educated in schools with instruction in English and non-proficient in Standard Arabic, English items are the only means to communicate in many fields. Such items become part of their mental lexicon and thus should be considered as borrowings even if they are not established loans.

Keywords: Arabic-English code-switching, borrowing, Egypt



1. Aim of the study

Code-switching, defined here after Li (2013: 360) as “a cover term to describe a range of linguistic behavior that involves the use of more than one language or language variety in the same interaction”, involves switching between both lone items and longer clusters. While multi-word insertions are unquestionable examples of code-switching, the status of lone items arouses controversy. In the literature on code-switching, lone items that are not established loanwords are in general treated as twofold, either as code-switches or nonce borrowings (for further explanation on the concepts see Section 3). The distinction between the two is usually based on grammatical factors.

The present study¹ deals with lone English items that occur in Arabic-English code-switching in the speech of students from the American University in Cairo. Our aim is to answer the question of whether such items are instances of code-switching or borrowing. In our data, nouns represent the most frequently switched category of lone items. At first glance, they can be classified as neither established loans, since they are not used by monolingual speakers, nor cultural loans that represent concepts and objects unfamiliar to Arabic speakers. However, a more detailed analysis, also taking into account sociolinguistic and pragmatic factors, reveals that their status is not so obvious. We claim that a lot of English items serve as established loans although they are not recognized by the whole Egyptian community. The reason is the linguistic situation in Egypt, the growing prestige of English which affects the education of the higher classes, and consequently the lack of proficiency in Standard Arabic among a specific sector of the Egyptian community.

2. The linguistic situation in the Arab world

2.1. Diglossia

The linguistic situation in the Arab world is usually referred to as diglossia. According to Ferguson's (1959) classical concept, diglossia implies that in a speech community two complementary varieties of a language coexist. These varieties fulfil different functions and are employed in different circumstances. The 'high' variety is primarily used for written purposes. As the language of literary heritage, it is highly standardized and enjoys a great deal of prestige. The 'low' variety, spoken in informal settings, is less prestigious and not standardized in terms of prescriptive grammars and dictionaries. In Arabic speech communities, diglossia includes Standard Arabic, unified throughout the whole Arab world, and a variety of vernaculars. Standard Arabic stands for the high variety. It is the language of high culture, literature, press and, importantly, the Koran and Islamic heritage which gives Standard Arabic the status of a sacred language within the Islamic community. Diverse vernaculars represent the 'low' variety. They are used in everyday communication and lack prestige attached to Standard Arabic. An important factor that sets Standard and vernacular Arabic apart from each other is their acquisition. Standard Arabic is acquired only through formal education. Vernaculars, on the other hand, are the

Arabs' first language, i.e. their mother tongue, irrespective of their class, level of education, religious affiliation or profession. Thus, proficiency in Standard Arabic, which translates into the ability to read and write, is only achieved by attending schools in which Standard Arabic is taught. In Egypt, the pursuit of Standard Arabic is gradually becoming weaker and weaker due to significant linguistic changes that arose from the growing interest in English and the emergence of bilingual code-switching.

2.2. English

Arabic-English code-switching in Egypt is a recent phenomenon. It emerged as a result of globalization and the open door policy initiated in the 1970s. Economic transformation paired with the partial privatization of the education system led to the growing popularity of private and international schools with instruction in English. Attending such schools that assured proficiency in English and consequently well-paid jobs in private companies (Schaub, 2000), became the common pursuit among the upper classes. English evolved into "the uncontested language of work and socializing in upper-middle class circles" (Koning, 2009: 61) and a hallmark of well-to-do Egyptians. Thus, mastery of English is a significant social marker supported by the strong stratification of the society.

The prevalence of English in the education of the upper classes has a great impact on the Egyptian linguistic landscape. In the majority of private and international schools with instruction in English, learning Arabic is not compulsory, or it is perceived as an unpleasant necessity (see Galegher, 2012). Their graduates, as Mehrez (2010: 210) points, constitute a young elite alienated from their own cultural heritage; "they all speak the vernacular, but for all intents and purposes they represent an illiterate elite where Arabic language and culture are concerned". As a result, English – as a social marker that affords a privileged position in the labor market – undermined the former position of Standard Arabic to become the most important linguistic capital in Egypt. Standard Arabic remains the domain of low-paid employees in the state administration and schooling (Haeri, 1997).

2.3. Code-switching

One of the consequences of the social prestige of English, its prevalence in the well-paid private sector and English-oriented education is the emergence of bilingual code-switching that is used as a form of in-group, private and professional, communication among the upper classes.

Nevertheless, code-switching is not a new phenomenon in Egypt. The sharp dichotomy between Standard Arabic and vernaculars has significantly weakened since the introduction of universal education in the 1950s and the dissemination of Standard Arabic nationwide. In actual language use, native speakers tend to combine elements from both varieties in a single speech event. This phenomenon is usually referred to as Educated Spoken Arabic (Mitchell, 1986), *'āmmiyyat al-muṭaqqafīn* (colloquial of the educated) (Badawi, 1973), mixed styles (Mejdell, 2006), multiglossia (Hary, 1996), or diglossic code-switching (Bassiouney, 2006; Boussofara-Omar, 2003).

Diglossic code-switching serves as a spoken equivalent of Standard Arabic. It is used in formal and semi-formal settings to discuss issues related to politics, culture, science, religion etc. Standard Arabic, as a primarily written variety, is perceived as too artificial to provide natural oral communication. Vernaculars, on the other hand, cannot fulfil all communicative needs, especially those related to more serious topics that go beyond the domain of everyday life due to, among others, the lack of the necessary vocabulary.

Thus, the linguistic situation in Egypt points to two things. Firstly, people who do not get a formal education in Standard Arabic – i.e. the majority of graduates of schools with instruction in foreign languages – are able to communicate in Arabic to a limited extent. Second, fields associated with the high variety in Arabic are accessible to them via English. Thus, a lot of English terms in fact became part of their mental lexicon – i.e. they are the only ones that are accessible during the language production process. This may indicate that this particular group among Egyptian society is not diglossic in fact (as far as the classical notion of diglossia is concerned). In their speech, the functions of the high and low variety are fulfilled by two unrelated linguistic codes – Egyptian Arabic and English.

3. Borrowing vs. code-switching

In the literature on code-switching, inserted lone items are treated as twofold. According to Myers-Scotton (1992; 1993; 2006) code-switching and borrowing are closely related processes. Therefore, single items should not be excluded from code-switching. Poplack and her associates (Budzhak-Jones and Poplack, 1997; Poplack and Meechan, 1998; Poplack et al., 1988; Poplack et al., 1989; Sankoff et al., 1990), on the other hand, make a strict distinction between borrowing and code-switching. Borrowing, as opposed to code-switching, involves the grammatical structure of one language. Thus, non-integrated lexical items are considered as nonce borrowings. These two different approaches to borrowing derive mainly from different theoretical approaches to code-switching developed by Poplack and Myers-Scotton.

3.1. Borrowing and code-switching as discrete processes

Poplack defines code-switching as “the juxtaposition of sentences or sentence fragments, each of which is internally consistent with the morphological and syntactic (and optionally, phonological) rules of the language of its provenance” (Poplack, 1993: 255). In other words, code-switching implies alternation between two linguistic codes. Switched elements and clusters observe the rules of the donor language throughout the discourse. Code-switching is controlled by two universal constraints. According to the free morpheme constraint, “codes may be switched after any constituent in discourse provided that constituent is not a bound morpheme” (Poplack, 1980: 585–86). The equivalence constraint states that code-switching is allowed as long as the juxtaposition of the elements from the two languages does not violate the syntactic rules of either language (Poplack, 1980: 586).

However, the universality of these constraints was questioned in many studies including those concerning code-switching with Arabic. Counterexamples of singly occurring lexical items were encountered in a variety of language pairs – Arabic and French (Bentahila and Davies, 1983); Arabic and Dutch (Nortier, 1990); as well as Arabic and English (Al-Khatib, 2003; Atawneh, 1992; Bader, 1995; Bader, 1998; Bader and Minnis, 2000; Mustafa and Al-Khatib, 1994). In Arabic-English code-switching, counterexamples to the free morpheme constraint include, among others, instances of mixed verbs composed of Arabic affixes and an English verb stem, e.g. *bitride* ‘she rides’ (Atawneh, 1992: 233) or nouns modified by the definite article *al/il*, a bound morpheme in Arabic. A typical violation of the equivalence constraint is exemplified by mixed adjectival phrases, e.g. *nose ṣḡīr* ‘a small nose’ (Atawneh, 1992: 230) where the word order follows the grammatical rules of Arabic and conflicts with the rules of English.

In response, Polack states that such cases are not instances of code-switching but borrowing. As opposed to code-switching, in borrowing only one linguistic system operates which results in the morphological and syntactical (though not necessarily phonological) integration of an inserted lexical item into the recipient language. Importantly, borrowing is not restrained to established loanwords incorporated into the lexicon of the recipient language and widely used by monolingual speakers. The use of non-established, non-phonologically integrated nor recurrent lexical items is termed *nonce borrowing* which refers to as “a one-off occurrence resorted to by the speaker” (Poplack et al., 1988: 58). Although *nonce borrowings* are usually not recognized by monolinguals, they are akin to established loanwords in terms of linguistic production. The so-defined borrowing and code-switching phenomena lead to the assumption that “lone major-class content words of one language incorporated in discourse of another are almost always borrowings” (Poplack and Meechan, 1998: 135). This hypothesis was tested in a variety of studies (Adalar and Tagliamonte, 1998; Budzhak-Jones, 1998; Eze, 1998; Turpin, 1998) that applied the comparative variationist methodology to examine syntactic context in which lone lexical items appear. These studies confirm that such items systematically behave as their counterparts in the other language and established loans. This entitles to conclude that they are cases of borrowing since only one linguistic structure is active at a time. Thus, according to this point of view, the only real cases of single-word codeswitches will be those which are not morphologically and syntactically incorporated into the recipient language.

3.2. Borrowing and code-switching as a continuum

Myers-Scotton, on the other hand, states that code-switching and borrowing are not different processes but part of the same continuum. This statement arises from an idea regarding the insertional nature of code-switching. According to Myers-Scotton’s Matrix Language Frame (MLF) model in bilingual production, there is always one dominant language at work and that is the matrix language (ML). The constraints on code-switching are based on the asymmetrical distribution between system and content morphemes. System morphemes, as it was originally postulated (Myers-Scotton, 1993), derive from the ML, whereas content morphemes may come from both the matrix and the embedded

language (EL). EL content morphemes are inserted into the ML, which provides the grammatical frame. Thus, non-attested in the ML EL lone lexical items are always framed by the ML since only the ML system operates.

The notion of continuum between borrowing and code-switching is based on the premise that code-switching takes an active part in introducing loanwords into the ML. However, it does not automatically categorize all lone lexical items inserted into the ML as cases of code-switching. Myers-Scotton makes a distinction between cultural loans, core borrowed forms and code-switching. Cultural loans are typical examples of borrowing. They refer to objects and concepts previously non-existent in the ML. Filling lexical gaps, they enter the ML lexicon abruptly and are widely used by both bilinguals and monolinguals who may not be aware of their foreign origins. As such, cultural loans, according to Myers-Scotton (1993: 173), ought to be excluded from code-switching as a phenomenon. Core loans, on the other hand, stand for objects and concepts already existing in the ML. Some of them may gradually become conventionalized code-switches due to their prestige and, eventually, part of the ML lexicon as established loans. Thus, code-switching is “the gate by which content morphemes as core B [borrowed] forms enter the ML” (Myers-Scotton, 1993: 174). To distinguish between code-switching and borrowing, Myers-Scotton proposes two criteria – predictability and frequency. Core borrowing forms, in contrast to code-switching forms, show a high frequency of occurrence and are relatively predictable; “It is not that a B [borrowed] form *must* recur, it is that CS [code-switched] form *must not* recur in order to be a CS form” (Myers-Scotton, 1992: 36, original emphasis). Specifically, content morphemes that appear at least three times in a relatively large corpus should be categorized as core borrowed forms (Myers-Scotton, 1993: 207).

Thus, both Poplack and Myers-Scotton utilize structural integrity to confirm the validity of the two models of code-switching. We agree with Myers-Scotton that code-switching and borrowing constitute a continuum. Not all single words of foreign origin automatically comprise the category of borrowings and therefore should be excluded from code-switching. Nevertheless, we believe that in Egypt not all non-established borrowings can be classified as belonging to code-switching. Secondly, the criterion here should be contextual/functional rather than purely structural; the local context as well as the speakers’ linguistic background should be included in the analysis.

3.3. Borrowing vs. code-switching in bilingual speech with Arabic

Most studies on code-switching with Arabic do not differentiate between the two phenomena, treating lone lexical items not established as loanwords in Arabic as code-switches (Al-Mansour, 1998; Atawneh, 1992; Bentahila and Davies, 1983; Bentahila and Davies, 1991; Boumans, 1998; Hussein and Shorrah, 1993; Mohamed, 1989; Myers-Scotton et al., 1996; Nortier, 1990; Sallo, 1994).

Mustafawi (2002) and Rouchdy (1992), on the other hand, make such a distinction, although in each case for different reasons. The first study takes a variationist approach to code-switching to test the nonce borrowing hypothesis. It investigates the morpho-syntactic behaviour of lone English-origin nouns focusing on patterns of determination,

gender assignment and word order – i.e. grammatical contexts incongruent in Arabic and English. The comparison of ambiguous lone nouns with established loans and native Arabic nouns in the contexts under investigation show that lone items are used in a similar way to Arabic nouns and established loans. This supports the nonce borrowing hypothesis and suggests, according to Mustafawi, that singly occurring items should be treated as borrowings.

A similar differentiation between borrowing and code-switching was adopted in Rouchdy (1992). However, in contrast to the purely syntactic approach of variationist studies, Rouchdy states that the use of code-switching vs. borrowing is determined by two factors – the linguistic system of the languages involved and the social context. A variable that seems to have a significant impact on the issue is the speaker's level of education. It determines both the degree of phonological adaptation and the patterns used. Semi-educated speakers tend to resort to borrowing whereas educated speakers use code-switching. Thus, the clause "She left her in the car" may be uttered as either *tarakitha bi-k-kār*, where *kār* is classified as a borrowed item, or *tarakitha in the car*, where the prepositional phrase *in the car* is defined as code-switching. The former would be typical of the semi-educated and the second of the educated. The same applies to pluralization patterns. Educated speakers avoid using the Arabic plural suffix *-āt* with English nouns, which is, on the other hand, the most common type of pluralization among semi-educated speakers (e.g. *parking lots* vs. *barkin lottāt*) (Rouchdy, 1992: 41). On the other hand, the semi-educated pronounce borrowed items closely to the English phonetics. In the speech of the educated, borrowed items are usually phonologically integrated with Arabic (Rouchdy, 1992: 42).

Thus, in both studies, the integration of inserted constituents is a decisive factor for the differentiation between code-switching and borrowing. Integrated items are *ad hoc* classified as instances of borrowing. However, Heath's study (1989) shows that in Moroccan Arabic there are items that function as loanwords, but show partial or no morphological integration. For instance, some French verbs introduced to Moroccan Arabic during the colonial period remained unintegrated in terms of inflection while more recent loans manifest some inflectional integration. This finding indicates different patterns of borrowing in Moroccan Arabic. Interestingly, in Moroccan Arabic-French code-switching, different patterns of code-switching are found as well. Bentahila and Davies' study (1992) on two generations of Moroccans reveals that balanced bilinguals – i.e. the older generation – tend to use inter-sentential code-switching, while the younger generation, with a lesser command of French, but more proficient in Standard Arabic, use intra-sentential code-switching. This may suggest that proficiency in both languages, as also suggested by Rouchdy, affects patterns of both code-switching and borrowing. Thus, we hypothesize that a lot of lone items that appear in the speech of fluent bilinguals in Egypt should be classified as instances of borrowing. However, this assumption is based not solely on bilinguals' linguistic proficiency. We claim that borrowing is also motivated by the linguistic situation in Egypt.

4. Data

The study is based on the data collected from students of the American University in Cairo in 2015. The AUC is the most prestigious and expensive university in Egypt with instruction in English. It attracts mostly fluent bilinguals from the upper classes graduated from private and international schools with instruction in foreign languages. Graduates of state schools who won scholarships at the university are subject to intensive language instruction after their admission. At the start of the study, students are supposed to have a fluent command of English.

For the present study, we used approximately 14 hours of recordings that consisted of eight group interviews designed to collect speech samples. The duration of the interviews ranged from 45 to 170 minutes. In each, at least two participants took part. The interviews took the form of relaxed conversation. To minimize ‘the observer’s paradox’ and get as much natural linguistic data as possible, we engaged the participants in conversation with each other. They discussed a variety of topics such as the linguistic situation in Egypt, code-switching, education, social inequalities, culture, religion, economy, sexual harassment, hobbies, family life, etc.

5. Methodology

The data were analyzed within the framework of the MLF model. Since the 1990s it has become the most frequently used theory in the study of code-switching. For the purpose of the present study, the MLF model is a more promising framework for analysis than the variationist approach. The latter consists of comparing the grammatical context of lone items from language 2 to that of language 1 and established loans. Thus, its application of the latter must end with the conclusion that all lone items are nonce borrowings since they almost always show morphological and syntactic integration with the recipient language. Otherwise the universality of the constraints formulated by Poplack will be questioned, which in fact was raised in a variety of studies on code-switching with Arabic.

The validity of the MLF model was questioned as well – e.g. in Aabi (1999) and Al-Enazi (2002). However, this is because those studies applied the MLF model in its original formulation, strictly adhering to the restrictions on system morphemes. As a result, in different data sets, system morphemes from the EL were detected, which is inconsistent with the main premise of the MLF model. Nevertheless, the asymmetry between the distribution of system and content morphemes does not assume that only content morphemes are allowed to be inserted into the ML. The Abstract Level and the 4-M model, developed as MLF sub-models (see e.g. Myers-Scotton and Jake, 2009), elaborate on morpheme classification with reference to their syntactic roles and distribution, admitting two types of system morphemes – early and bridge – to occur in bilingual constituents. Studies that incorporate these amendments into the analysis (Alenezi, 2006; Al-Rowais, 2012; Okasha, 1999) confirm the validity of the MLF model in bilingual code-switching with Arabic. Furthermore, the MLF model does not automatically exclude integrated items from code-switching. Given that the MLF model

is insertional, both code-switches and borrowed items should be subject to morphological and syntactic integration.

The data were divided into projections of complementizer (CP) – the units of analysis in the MLF model. The individual CPs were further divided into four main categories: (1) Arabic monolingual CPs, (2) CPs with Arabic as the ML, (3) English monolingual CPs, and CPs with English as the ML. Any CPs with composite ML – i.e. CPs in which both languages simultaneously provide the grammatical frame (see e.g. Myers-Scotton, 1998) – were excluded from the analysis.² Lone items in bilingual CPs were subject to further analysis which included both singly occurring and recurrent ones. We did not exclude recurrent items (which Myers-Scotton classifies as core borrowed forms) since a lot of them seemed to occur as a consequence of linguistic accommodation to the interlocutor. Lone items were categorized according to the type of morphemes they represent. The following categories were distinguished: adjectives, adverbs, conjunctions, definite articles, discourse markers, nouns, personal pronouns, prepositions and verbs. The goal of this procedure was to find any possible differences between the two types of CPs since they may translate into functional differences between the two languages. The analysis also includes recurrent compound nouns. In the MLF model inserted constituents are not necessarily single words. A constituent can be any syntactic unit, either a lexical item (e.g. a noun) or a phrase. Constituents of more than one word are called ‘EL islands’. Such islands are well-formed according to the grammatical rules of the EL. Within this approach, compound nouns should be treated as EL islands. However, their recurrence and their lack of Arabic equivalences in the data indicate that they are perceived as single units. For that reason, we differentiate between this particular kind of compound noun and the ‘real’ EL islands that were investigated at the final stage of the study.

6. Results of the study

Table 1 shows the distribution of the categories found in particular interviews. AR M stands for monolingual Arabic CPs, AR ML – CPs with Arabic as the ML, ENG M – monolingual English CPs, and ENG ML – CPs with English as the ML. The data differ in the total number of CPs ranging from 865 in interview 1 to 3575 in interview 6 due to differences in the duration of the interviews. They are also significantly diversified in terms of patterns used. Interview 8 exhibits the smallest number of CPs with English providing the grammatical frame, namely seven CPs (five monolingual and two mixed) which constitutes 0.73 percent of the total CP number. The highest number of CPs controlled by English is found in interview 3, with 596 CPs (56.65 percent). Nevertheless, the quantitative analysis of the whole data shows two features common to all data sets. The first one is the participants’ preference for monolingual CPs. The second one is their preference for Arabic over English CPs within monolingual CPs. Out of the total number of 13513 CPs, 10170 CPs (75.26%) are either monolingual Arabic (6578) or monolingual English (3412). Mixed CPs that were subject to analysis number 3443 (2076 with Arabic

and 1367 with English as the ML) which constitutes approximately 25 percent (15.36% and 10.12% for Arabic and English respectively).

		total CP	AR M	AR ML	ENG M	ENG ML
interview 1	N	865	556	185	94	30
	%		64,28%	21,39%	10,87%	3,47%
interview 2	N	928	402	121	287	118
	%		43,32%	13,04%	30,93%	12,72%
interview 3	N	1052	350	106	391	205
	%		33,27%	10,08%	37,17%	19,49%
interview 4	N	1264	885	244	76	59
	%		70,02%	19,30%	6,01%	4,67%
interview 5	N	2047	1002	281	607	157
	%		48,95%	13,73%	29,65%	7,67%
interview 6	N	3575	1757	611	706	501
	%		49,15%	17,09%	19,75%	14,01%
interview 7	N	2825	1084	300	1146	295
	%		38,37%	10,62%	40,57%	10,44%
interview 8	N	957	722	228	5	2
	%		75,44%	23,82%	0,52%	0,21%
Total	N	13513	6758	2076	3412	1367
	%		50,01%	15,36%	25,25%	10,12%

Table 1: Overall distribution of monolingual and bilingual CPs

Single constituents in mixed CPs with Arabic as the ML (including compound nouns) number 1843. In CPs with English as the ML, single constituents amount to 1116. Tables 2 and 3 display the type and frequency of constituents that constitute at least 0.5% of all constituents in Arabic and English CPs respectively in order of their frequency of occurrence. The distribution of particular types of morphemes differs significantly. The most frequently inserted morphemes are nouns followed by adjectives, verbs and adverbs in CPs with Arabic as the ML, and conjunctions, discourse markers, adverbs and nouns in CPs with English as the ML.

	N	%
nouns	1318	71,51%
compound nouns	164	8,90%
adjectives	162	8,79%
verbs	68	3,69%
adverbs	61	3,31%
conjunctions	29	1,57%
prepositions	15	0,81%
Total	1817	98,58

Table 2: Mixed CPs with Arabic as the ML

	N	%
conjunctions	437	39,16%
discourse markers	314	28,14%
adverbs	196	17,56%
nouns	87	7,80%
personal pronouns	37	3,32%
adjectives	12	1,08%
prepositions	12	1,08%
definite articles	7	0,63%
Total	1102	97,69

Table 3: Mixed CPs with English as the ML

6.1. Mixed CPs with Arabic as the ML

Nouns constitute 71.5 percent of the total number of single-constituent insertions in CPs with Arabic as the ML. If we add compound nouns, the percentage increases to 80.4 percent. The majority of the inserted nouns are related to academic life, professional career, social and political issues, culture as well as geographical names (e.g. course, student, dorm, professor, English, media, sexual harassment, social class, high school

etc.) In general, these are terms that are acquired through formal education and associated with Standard Arabic in monolingual discourse. Most of them, whether single or compound nouns, are used as bare forms fully integrated morphologically and syntactically with Arabic, which is exemplified in (1) and (2) below. In (1), *program* lacks an indefinite article which is absent in Arabic. In (2), *middle class* is modified by the Arabic article *il-*, the verb *iḥtafa* ‘to disappear’ is used in the feminine form. The latter indicates that the speaker knows the Arabic equivalent of *class* – *ṭabaqa* – which is feminine. However, the knowledge of Arab equivalents is not obvious. It can be tested solely when the noun is modified by a determiner, adjective, verb or in possessive constructions with *bitā* ‘of’ – which is subject to declination according to gender and number – provided that the Arabic equivalent is feminine. In such contexts, complete syntactic integration is not always observed. In (3) *economy* is used with the feminine demonstrative *di* and the feminine form of the verb *assar* ‘to affect’ although the Arabic equivalent *iqtiṣād* is masculine. In (4) *graduation project*, which in Arabic is masculine (*mašrū* ‘*taḥarruġ*’), is followed by the feminine form of *bitā* ‘.

- (1) kunn iḥna mašyīn taba‘ **program** mu‘ayyan
 were.1PL we going accordance program specific
 We were proceeding according to a specific program.
- (2) il-**middle class** ‘and-ina iḥtaf-it ta’rīban
 the-middle class with-our disappeared-PRFX.3f approximately
 Our middle class has almost disappeared.
- (3) fa b-adris izzāy ba’a l-**economy** di bi-t’assar ‘a n-nās
 so ASP-learn.1s how so the-economy this.FEM ASP-affect.3fs on the-people
 So I learn how this economy affects people.
- (4) ‘and-i **graduation project** bitā‘it **con simulation**
 with-my graduation project of.FEM.SG con simulation
 I’m doing a graduation project on CON simulation.

The integration of nouns with Arabic is absent in plural forms and infrequently observed in forms modified by the article *il-* where its assimilation to the first phoneme of the noun is required. Plural forms, both single and compound nouns, are inserted with the suffix *-s* (5):

- (5) ana kull illi ‘and-i **numbers** u **consumer trends**
 I all that with-me numbers and consumer trends
 All that I have is numbers and consumer trends.

There is one exception. The noun *courses* is sometimes pluralized with the Arabic suffix *-āt*. Interestingly, both forms may be used by the same speaker, as it is in (6) and (7):

- (6) b-yaḥdu kulli l-ḥagāt wi-l-**courses** aw il-**classes** bita‘t-um bi-l-**English**

ASP-they take all the-things and the-courses or the-classes of.FEM-their in-the-English
They take all things and courses or their classes in English.

- (7) mu‘ḏam il-**kurs-āt** bi-l-**English**
majority the-course-SFX.PL in-the-English
The majority of the courses are in English.

The phonological assimilation of *il-* involves so called sun letters (*t, ṭ, d, ḏ, r, z, s, š, ṣ, ḍ, ṡ, ḏ, l, n*). Whenever the article is followed by one of these consonants the *l* in the article assimilates to it which results in a doubled consonant. Although the assimilation (8) appears on a regular basis, there are quite a few instances where it is not observed (9, 10). The reason for not adhering to the assimilation rules is not obvious. It is certainly determined by neither particular items nor consonants. Even high frequently used items, such as *social class* or *topic*, appear in the speech of some participants with unassimilated articles. This may indicate that some speakers prefer not to assimilate the article to English nouns. However, this assumption requires a different research design to allow its verification.

- (8) wi ma-fī-š **contradiction** ma bēn is-**science** wi r-**religion**
and NEG-there_is-NEG contradiction PART between the-science and the-religion
There is no contradiction between science and religion.
- (9) bass il-**struggle**ḥaṣal lamma gīt AUC
but the-struggle happened.3ms when came.1s AUC
The struggle happened when I came to AUC.
- (10) izzāy il-filūs bi-tḥuṣṣi u tiṭla‘ (.) il-**stock market**
how the-money ASP-enters.z3f and go-out.3f the-stock market
How money goes into and out of the stock market.

Adjectives are inserted as bare forms not observing the gender or number congruence with the noun when required in Arabic. They usually act as predicates, which is exemplified in (11) and (12). In both CPs, Arabic requires adjective inflection for the plural in (11) and feminine in (12). When used as attributives, adjectives follow the noun adhering to the Arabic word order (13). They are also modified by the definite article (14) if it is obligatory in Arabic.

- (11) iḥna miš ‘ayzīn yib’u **cosmopolitan**
we NEG modal_of_desire.PL are.3 cosmopolitan
We do not want to be cosmopolitan.
- (12) fa hiyya **rural** ’awi
so she rural very
So it’s very rural.
- (13) aw yib’a fih **hina**’āt **physical** ma bēn rāgil u sitt
or is.3MASC there_is fights physical PART between man and woman
Or there are physical fights between a man and a woman.

- (14) ana b-aḥibb iṣ-ša‘r is-**straight** wi ṣ-ša‘r il-fātiḥ
 I ASP-like the-hair the-straight and the-hair the-fair
 I like the straight hair and the fair hair.

Out of 68 verbs recorded in the data 56 take an Arabic inflection (15). The remaining 12 are inserted as bare forms. Integration with the Arabic conjugation or its lack seems to depend on the grammatical context in which a verb occurs. For instance, the third plural forms are inflected with the prefix, but not with the suffix *-u* marking a plural. The difference between Arabic and English verbs in the plural is apparent in (16) – the verb *accommodate* is inflected only with the prefix *yi-* while the preceding verb *byiḥawlu* ‘they try’ with both the prefix *yi-* and the suffix *-u*. However, the plural suffix appears with English nouns that are used as verbs (17).

- (15) ana kamān ha-a-**accept** il-mawḍū‘ da
 I also FUT-PFX.1s-accept the-matter this
 I’ll also accept this matter.
- (16) inn **professors** b-yi-ḥawl-u yi-**accommodate** li **everyone**
 that professors ASP-PFX-try-SFX.3PL PFX-accommodate to everyone
 That professors try to accommodate to everyone.
- (17) wi ba’ū b-yi-**target**-u il-**artists**
 and began.3PL ASP-PFX.3ms-target-SFX.3PL the-artists
 And they began to target the artists.

6.2. Mixed CPs with English as the ML

As opposed to CPs with Arabic as the ML, in CPs with English as the ML nouns come only fourth. However, the relatively high position of nouns is due to interview 1 in which 53 nouns out of 87 appear mostly in one participant’s utterances. A total of 51 of them refer to objects and concepts related to Arabic and Islamic culture. They seem to be used for two reasons. First, often their English equivalents may connote slightly different meanings, e.g. *musalsalāt*, which stands for TV shows produced for Ramadan, is usually translated as ‘soap operas’ although the terms refer to distinct socio-cultural phenomena. The second reason seems to be more of religious nature. Religious discourse on Islam in the Arab world is believed to be delivered in Standard Arabic. In the interview in question, one of the participants discussed religious issues primarily in English due to, as she admitted, her insufficient knowledge of Standard Arabic. The use of Arabic religious terms was a way to authenticate the message and emphasize her religious devotion. Such terms are introduced as either lone items or EL islands, e.g. *āyat il-kursi* (The Throne Verse), *sayyidina muḥammad* (our lord Muhammad). While inserted as lone items they are mostly morphologically and syntactically integrated to English which includes, for instance, their modification by the articles *a* (18) and *the* (19) or the plural suffix *-s* (20):

- (18) **today they tell you that a woman cannot be part of the mosque, cannot be an** imām (imam), **cannot be a** šēḥ-a (scholar trained in the religious sciences-FEM)
- (19) **the** šēḥ **was very angry with me**
- (20) **my mother has always been partial to the** maḡribi-s (Morrocans)

In the remaining interviews, the number of nouns ranges from zero to ten with an average of approximately five nouns per interview. The low number of lone nouns goes hand in hand with an even lower number of adjectives (1.08%) and a complete lack of inserted verbs – i.e. the most frequent categories in CPs with Arabic as the ML.

The most numerous categories in CPs with English as the matrix language are conjunctions (39.16%), discourse markers (28.14%) and adverbs (17.56%). The difference in the two types of CPs lies not only in the categories used, but also in the patterns of code-switching. In CPs with English as the ML, code-switching is more conventionalized and less insertional than in CPs with Arabic as the ML. By ‘conventionalized’ I mean that it is more predictable in terms of the prevalence of specific items. The most recurrent conjunctions are *fa* ‘so’ (21) and *bass* ‘but’ (22). In the category of discourse markers *ya‘ni* ‘it means’ (23), which is typical of Egyptian Arabic native speakers, prevail.

- (21) *fa* (so) **a lot of them can’t find jobs**
- (22) *bass* (but) **you don’t have an expire date**
- (23) **couse** *ya‘ni* (I mean) **I was so used to like this crazy life over there**

Adverbs show greater variety compared to conjunctions and discourse markers. Their use seems to be related to the participants’ preferences. However, the majority of them display a high frequency of occurrence in Egyptian Arabic, e.g. *barḏu* ‘also’ (24), *ḥālīṣ* ‘at all’ (25), *aṣlan* ‘basically’ (26) etc.

- (24) *bass* (but) **it’s a matter of choice** *barḏu* (as well)
- (25) *ya‘ni* (I mean) **they weren’t accepting that** *ḥālīṣ* (at all)
- (26) **which is** *aṣlan* (basically) **very expensive**

Importantly, such items are not grammatically dependent elements. In contrast to nouns, adjectives and verbs, their use is not restricted by the lack of congruency between Arabic and English. Thus, they are not genuine insertions but rather represent a pattern of intra-sentential code-switching that is referred to by Muysken (1997; 2000; 2007) as alternational code-mixing.

Muysken distinguishes three types of code-mixing: insertion, alternation, and congruent lexicalization. Insertional code-mixing resembles the MLF model – i.e. a single constituent, be it a lexical item or a phrase, from language 1 is inserted into the

grammatical structure of language 2. It is characterized by a nested *a b a* structure where ‘a’ stands for the ML and ‘b’ the EL. Inserted constituents are usually nouns and adjectives.

In alternation, on the other hand, there is no ML. This is when one language switches to the other without a structural relationship between the switched clusters that are usually longer than insertions and consist of several constituents. Alternation usually happens at the major clause boundary and involves, among others, conjunctions, adverbial modification or discourse marker switching which in Muysken (2007) was distinguished as a fourth type of code-mixing.

All the examples cited above (21-26) meet the conditions to be classified as alternations. This means that CPs with Arabic as the ML and CPs with English as the ML are different types of code-switching. This will be even more obvious if we look at multiword alternations that were initially classified according to the MLF model as EL islands.

6.3. Embedded language islands

6.3.1. CPs with Arabic as the ML

CPs with EL islands constitute 24.56 percent (600) of the total number of CPs with Arabic as the ML and 26.10 percent (396) of CPs with English as the ML. In CPs with Arabic as the ML, most EL islands are noun phrases. They are well-defined according to the rules of English – i.e. when modified by adjectives they follow English word order (27). Such phrases may occur with the articles *a/an* (28) and *the* (29). However, they are used inconsistently, as can be seen in (29) where the same phrase is repeated twice with *the* and then occurs with the Arabic *il-*.

- (27) wi sa‘āt law inta bi-ti-kallim bi l-**proper accent**
and sometimes if you ASP-PFX.2SG-speak with the-proper accent
And sometimes when you speak with the proper accent.
- (28) fa fi ṣ-ṣīn ḥatta gabū-l-na **an Arabic tutor**
so in the-China even brought.3PL-to-us an Arabic tutor
So in China, they even brought us an Arabic tutor.
- (29) ‘arfīn **this main street?** ‘arfīn **this main street?** il-**main street** illi barra da?
knowing.PL this main street, knowing.PL this main street, the-main street that outside this?
Do you know this main street? Do you know this main street? The main street that is outside?

As for the article *a*, it is barely used since Arabic lacks the indefinite article. Hence, the majority of noun phrases are inserted without the article, observing the rules of Arabic (30). However, the indefinite article is always present if a noun is preceded by the quantifier *a lot* (31):

- (30) fih **Egyptian community** kbīr f holanda
 there_is Egyptian community big in the_Netherlands
 There is a big Egyptian community in the Netherlands.
- (31) **a lot of the professors** fa‘lan b-yi-nazzil-u l-level
 a lot of the professors really ASP-PFX.3-reduce-SFX.3p the-level
 A lot of the professors really reduce the level.

EL islands are usually nested – i.e. they are inserted into the grammatical structure of Arabic. In (30) the word order of the phrase *Egyptian community* is English, although as a part of a larger phrase ‘a big Egyptian community’ is followed by the adjective *kbīr* ‘big’ according to the rules of Arabic since Arabic is the ML of the whole CP.

Two types of English modifiers are always used in EL islands – possessive determiners (32) and numerals (33). This probably arises from the incongruence between Arabic and English. Possessive determiners in Arabic are enclitics which never appear with English items (similarly to suffixes in English verb inflection). Numerals in Arabic have a complex syntax – nouns that follow numerals may be used either in singular or plural depending on the numeral.

- (32) ana masalan **my friends** miš b-yi-kallim-u **English** ktīr
 I for_example my friends NEG ASP-PFX-speak-SFX.3PL English much
 For example, my friends don’t speak English a lot.
- (33) ruḥna bi **two minibuses**
 went.2PL by two minibuses
 We went by two minibuses.

6.3.2. CPs with English as ML

In CPs with English as the ML, most EL islands are prepositional phrases (34) usually used in adverbial function and frozen expressions of high frequency of occurrence in Egyptian Arabic.

- (34) fa (so) **we stopped having classes together** ba‘di sana u-nuṣṣ (after a year and a half)

However, most Arabic clusters that occur in English CPs are difficult to classify as ‘real’ insertions since they do not form single constituents – i.e. the elements that compose them are not syntactically related to each other. The majority of such clusters are combinations of three elements: adverbs, discourse markers and conjunctions. In (35) the cluster is composed of the conjunction *bass* ‘but’ and the discourse marker *ya ‘ni*. In (36) the conjunction *lākin* ‘but’ with the adverb *dilwa’ti* ‘now’ occurs.

- (35) bass ya ‘ni (but I mean) **it takes me a lot of efforts**
- (36) lākin dilwa’ti (but now) **in our age a lot of people stop**

The second most frequently occurring Arabic clusters are those that include personal pronouns:

(37) inn ana (that I) **I'm betrayed as a friend**

(38) ya'ni ihna (I mean we) **we invest a lot in weddings** ya'ni (I mean)

(39) wa ya'ni humma (and I mean they) **they can't afford**

Personal pronouns mostly appear immediately before their English equivalents, resulting in pronoun doubling. The phenomenon seems to be typical of Arabic bilingual code-switching since it was observed in data from code-switching with French (Bentahila and Davies, 1983), Dutch (Nortier, 1990) as well as English (Eid, 1992; Okasha, 1999). Arabic personal pronouns before English pronouns also occur as lone elements (40).

(40) ihna (we) **we're renting the house**

Myers-Scotton and her collaborators (Jake, 1994; Myers-Scotton, 2010; Myers-Scotton et al., 1996; Okasha, 1999) explain pronoun doubling as a consequence of the incongruence in the pronoun system of Arabic and English that serves as a predictable way to satisfy the requirements of English and Arabic. Muysken (2000: 181), on the other hand, states that pronoun doubling is an instance of alternation since the doubled pronoun appears in the left-dislocated position. Left-dislocation (not resulting in pronoun doubling) is very common in monolingual Egyptian. When a dislocated element is a pronoun it may precede, for example, a noun or prepositional phrase. A similar pattern is found in (41) and (42). In CPs with English as the ML left-dislocation includes not only pronouns. In (43), the phrase *btū' il-handasa* is referred to again as *they*. In (44) the dislocated element *ahli* 'my family' appears in a longer monolingual cluster. Since switching between a noun and a verb, as can be found in (45), is possible, we state that left-dislocation (of nouns as well as pronouns) primarily serves pragmatic functions.

(41) ana (I) **my school was expensive**

(42) ya'ni ihna **for us he would have to be a Copt like me**

(43) btū' il-handasa (of.PL the-engineering, i.e. those in engineering) **they used to speak more Arabic**

(44) bass ana ahli masalan (but I my family for example), **they don't pray as much**

(45) **andnow** il-banāt (the girls) **are the majority**

7. Discussion

The analysis thus far shows a clear discrepancy between the categories of items used in the two types of mixed CPs. In CPs with Arabic as the ML, the most commonly switched category is English nouns followed by adjectives and verbs. These categories, as open class words that easily accept new members, are the most frequently borrowed items. In CPs with English as the ML, the majority of switches are conjunctions, discourse markers followed by adverbs, nouns and personal pronouns. Thus, the open class words most numerous represented here are adverbs. Nouns, which clearly outnumber the other categories in CPs with Arabic as the ML, occur after conjunctions, discourse markers and adverbs. However, as mentioned, the relatively high number of nouns in CPs with English as the ML is due to one interview in which the majority of the recorded nouns appear. Adjectives constitute only one percent, while Arabic verbs do not occur at all. Therefore, with the exception of discourse markers and adverbs, closed class words prevail. The analysis of multiword-switches shows a similar tendency. English EL islands centre around nouns – i.e. they are mostly noun phrases. In CPs with English as the ML, the majority of multiword-switches are adverbial phrases and clusters consisting of conjunctions, discourse markers, adverbs and personal pronouns.

If we look only at the grammatical aspect of the results obtained, we have to admit that in general they corroborate both Myers-Scotton's and (to a lesser degree) Poplack's approach to borrowing as far as CPs with Arabic as the ML are concerned. If our point of reference is the nonce borrowing hypothesis, we should recognize most singly occurring items as instances of nonce borrowing since the majority are morphologically and syntactically integrated with Arabic. The phonological assimilation of the definite article *il-* to the subsequent item, which occurs on a regular basis, further supports this statement. This, however, does not include plurals, mostly inserted with the English suffix *-s*, as well as adjectives that occur as bare forms. The lack of morphological integration suggests that such items should be classified as code-switches which, on the other hand, contradicts the idea that most lone items from language 2 are nonce borrowings. In the data, there is only one item that happens to be pluralized with the Arabic *-āt*, i.e. *kursāt* 'courses'. However, it is used by the same participants interchangeably with its English equivalent. Thus, English plural nouns and adjectives inserted into Arabic seem to be the main challenge for the nonce borrowing hypothesis. Surprisingly, in Mustafawi's (2002) study, which is solely designed to test the hypothesis, the issue is omitted.

The morpho-syntactic integration is the reason for which the same insertions will be classified in the MLF model as either code-switches, conventionalized code-switches or core borrowed forms in the case of frequently recurrent and predictable items. Again, a problem occurs with plural items. A lot of them – e.g. *classes*, *courses* – occur recurrently and their use is predictable to a high degree (although their Arabic equivalents may appear as well). This indicates that they should be classified as conventionalized code-switches. The plural suffix *-s* is categorized under the 4M model as belonging to early system morphemes that are indirectly selected by content morphemes to specify their meaning (see e.g. Jake and Myers-Scotton, 2009). Thus, its occurrence in CPs with Arabic as the

ML does not contradict the main premise of the MLF model – that of the asymmetrical distribution of content and system morphemes.

However, a more detailed analysis of lone items as well as compound nouns, indicates that in fact they are borrowings. This is not to say that they are widely accepted by monolingual speakers. On the contrary, they are comprehensible only by fluent bilinguals becoming, however, part of their mental lexicon. In the speech of those speakers they function as borrowed items mainly due to their bilingual education paired with the linguistic situation in Egypt.

The majority of inserted nouns relate broadly to the academic life or fields associated with Standard Arabic in monolingual discourse. The structural context in which they appear shows that they do not always fill lexical gaps since the participants seem to know their Arabic equivalents. This may suggest that inserted nouns are indeed code-switches. However, there are still a lot of instances that lack gender congruence indicating the lack of their Arabic equivalents at the speakers' disposal.

According to the participants' testimonies, they use English items, compound nouns as well as expressions (e.g. *have a nice day*) for two reasons. In a lot of circumstances, they are the only forms accessible to them due to their education in English and exposure to American popular culture since their childhood. Lewko (2012) found similar statements in his interviews with AUC students which confirms the universality of the phenomenon in the AUC community. Some participants in the present study even claimed that they had to use English to convey specific meanings since Arabic did not provide them (even if in reality they do exist in Arabic). The second reason is that some Arabic items/expressions are perceived as inappropriate since they do not cover the exact meaning connoted by their English equivalents. This also includes items that seem to have exact equivalents in Arabic – e.g. anthropology.

For the majority of the participants, raised in and accustomed to the two languages, code-switching is a natural way of communicating with others. The participants stressed that they do so subconsciously:

“It’s very natural, I mean, for the first time I paid attention that I spoke English and Arabic in the same sentence when foreigners asked me, ‘Why are you speaking like that?’ For me it’s a need [...] when I meet somebody I automatically know which language to speak, I don’t think, I mean, the process for me is very natural.”

Speaking in monolingual Arabic as well as monolingual English is certainly an uncomfortable challenge. One participant explained his experiences while dealing with monolingual Arabic speakers as follows:

“I have to translate it in my head, so the best, like I’m most comfortable and most confident with people who speak both [Arabic and English] because I can then like navigate freely but otherwise when I feel like I’m constrained by just one language.”

Such statements paired with the analysis of the participants' linguistic behaviour (in terms of grammar as well as pragmatics) indicate that both Arabic and English are part

of their mental lexicon. Due to their bilingual education and exposure to two different cultures they are aware of the cognitive differences between the two cultures and the languages that relate the cultures. English is often used to convey messages related to more serious topics and to do so in a precise and effortless way. Moreover, their type of formal education makes them unfamiliar with Arabic items that belong to the domain of Standard Arabic although they are fluent in vernacular. Thus, the majority of items/compound nouns that occur in CPs with Arabic as the ML are in fact borrowings since in reality they fill lexical gaps as well cultural ones that relate to the two languages.

We claim that this also includes items pluralized with *-s*. Rouchdy (1992) states that educated speakers tend to use English plural items with *-s* as opposed to the semi-educated who resort to the Arabic suffix *-āt*. The author recognizes morphological integration of plurals as a decisive factor in the differentiation between code-switching and borrowing and hence classifies the first phenomenon as code-switching and the second one as borrowing. The term 'educated' refers to communicative bilinguals in both languages. Therefore, the linguistic profile of the participants in the present study more or less corresponds with that of the participants in Rouchdy's study. This may indicate that English plurals are typical of fluent bilinguals regardless of whether they are code-switches or instances of borrowing.

The discrepancy between the two types of CPs further supports the claim that the majority of lone English items are borrowings. Arabic lone morphemes are functional items (conjunction, pronouns), discourse markers and adverbs. Typical EL islands – i.e. those that show an internal structural relationship – form the minority. Longer Arabic clusters are composed of function words, discourse markers and adverbs. As such, they are not insertions but alternations. According to Muysken, this type of code-switching is especially frequent in stable bilingual communities with a tradition of language separation while insertion is typical of communities in which speakers are not equally proficient in the two languages – e.g. recent migrant communities. A similar phenomenon is also found in the AUC community. AUC students do not constitute a monolithic community in terms of their education and backgrounds as is often claimed. Besides those educated in schools with foreign languages in Egypt, there are also students who lived abroad most of their lives as well as those who went to public schools. The type of exposure to English, their linguistic behaviour outside the AUC and so on clearly affects the patterns of code-switching they use. Insertional code-switching with Arabic as the ML is the only one that appears in the speech of all students including graduates of public schools. Alternation, on the other hand, is typical of bilinguals with long-term exposure to English and almost the only one type of code-switching used by those who feel more comfortable speaking English.

These differences in patterns of switching between Arabic and English may indicate that the two languages are used for specific pragmatic reasons. While insertional code-switching is used primarily for lexical references, alternation allows the discourse to be adapted to the local environment in a relatively easy way. Such embedding is desirable for a variety of reasons – for example, while discussing religious issues that are traditionally associated with Standard Arabic. More detailed pragmatic analysis, which

is beyond the scope of the study, shows further evidence for the distributional and functional asymmetry between Arabic and English which argues for the claim that English is taking over particular fields attributed to Standard Arabic in monolingual discourse. Thus, we should conclude that the AUC community is diglossic in a broader sense of diglossia (so-called extended diglossia, see Fishman, 1967) where the two varieties of a language are replaced by two unrelated languages – i.e. Egyptian Arabic and English. Since English supplies mainly the lexical needs that arise from bilinguals' socio-linguistic and cultural background, the majority of English lone insertions in Arabic should be acknowledged as borrowed rather than code-switched items even though they are not established loans.

8. Conclusion

The study aimed to determine whether lone English items in Arabic-English code-switching are borrowed forms or code-switches. Our hypothesis was that the majority of such items would constitute borrowed forms due to the linguistic situation in Egypt. The hypothesis was tested within the framework of the MLF model. Two types of bilingual CPs were investigated – CPs with Arabic as the ML and CPs with English as the ML. A comparative analysis demonstrated that the two types of CPs differ in the categories of morphemes used. English provides nouns, adjectives and verbs. The majority of them relate to fields that are associated with Standard Arabic in monolingual discourse. According to the participants in the study, their use is determined by the speakers' insufficient knowledge of Standard Arabic due to their education in schools with instruction in English and their life-long exposure to Western culture and media. Thus, English items primarily fulfil the lexical needs to supplement the bilinguals' vocabulary and, hence, should be classified as borrowed items. In previous studies, the morphological integration to the dominant language was seen as a pivotal factor in the differentiation between code-switching and borrowing. The results of the study show that it is not always preserved in the data. This includes plural items, adjectives and, to some extent, verbs. These categories are either suffixed with the English -s (plural nouns), used as bare forms without a feminine marker (adjective), or partially inflected by Arabic prefixes but not suffixes (verbs). We state that the lack of morphological integration results from the incongruence between the two languages. The participants in the study are proficient bilinguals, aware of the structural differences between Arabic and English. This, in turn, translates into a partial integration of borrowed items. Therefore, we believe that the analysis of code-switching vs. borrowing should incorporate the structural characteristics of the languages involved, functional dimension, the speakers' sociolinguistic and psycholinguistic background, and the local linguistic situation.

Notes

1. The research was financed by the National Science Center in Poland based on decision number DEC-2013/11/D/HS2/04524. The data were collected and transcribed with co-researcher Magdalena Zawrotna.

2. In the MLF model, mixed CPs must conform to two fundamental principles – the Morpheme Order Principle and the System Morpheme Principle. The first one defines the ML as the only one that supplies the word order of a CP. The System Morpheme Principle (supported by the 4-M model) prohibits the occurrence of EL outsider system morphemes, typically subject-verb agreement and case-marking, that “coindex relations that hold across phrase and clause boundaries” (Jake and Myers-Scotton 2009: 225). If either of these two principles is not satisfied, neither of the languages can be defined as the ML; the result is a composite ML with two languages providing the grammatical structure. This is the case in the example below. The CP exemplifies a typical Arabic nominal sentence expressing possession with Arabic supplying all outsider system morphemes. However, the word order of the phrase ‘*arabi classes*’ observes the rules of English, which violates the Morpheme Order Principle. Thus, the CP is simultaneously framed by the two languages. We decided to exclude such CPs for the sake of the clarity of the analysis, especially that they are few in number in the data.

kān ‘and-ina ‘arabi **classes**

was.1MASC with-our Arabic.MSC.SG classes

We had Arabic classes.

References

- Aabi, Mustapha (1999): *The Syntax of Moroccan Arabic/French and Moroccan Arabic/Standard Arabic Code Switching*. PhD dissertation. University of Sheffield.
- Adalar, Nevin, and Sali Tagliamonte (1998): “Borrowed nouns; bilingual people: The case of the ‘Londrali’ in Northern Cyprus”. *International Journal of Bilingualism*, 2(2): 139-159.
- Al-Enazi, Mohamed Husein (2002): *The Syntactic Form and Social Functions of Saudi Arabic-English Code-Switching among Bilingual Saudis in the United States*. PhD dissertation. Indiana University of Pennsylvania.
- Alenezi, Farah H. (2006): *Formal Constraints on Arabic/English Code-Switching: A Lexically-Based Approach*. PhD dissertation. University of Kansas.
- Al-Khatib, Hayat (2003): “Language alternation among Arabic and English youth bilinguals: Reflecting or constructing social realities?” *International Journal of Bilingual Education and Bilingualism*, 6(6): 409-422.
- Al-Mansour, Nasser Saleh (1998): *Linguistic Constraints on Code-Switching: A Case Study of Saudi- Spoken Arabic English Code-Switching*. PhD dissertation. University of Florida.
- Al-Rowais, Hawazen. (2012): *Code Switching between Arabic and English, Social Motivations and Structural Constraints*. PhD dissertation. Ball State University.
- Atawneh, Ahmad (1992): “Codemixing in Arabic-English bilinguals”. In E. Broselow et al., eds., *Perspectives on Arabic Linguistics IV: Papers from the Fourth Annual Symposium on Arabic Linguistics*. Amsterdam/Philadelphia: John Benjamins Publishing, 219-241.
- Badawi, El-Said M. (1973): *Mustawayāt Al-Luġa Al-‘arabiyya Al-Mu‘aṣira Fī Miṣr*. PhD dissertation. Dār al-ma‘ārif.

- Bader, Yousef (1995): "Code-Switching to English in daily conversations in Jordan: factors and attitudes". *Abhath Al-Yarmouk*, 13(2): 9–27.
- Bader, Yousef (1998): "Lexical codeswitching in the speech of an Arabic- English bilingual child". *Interface*, 13(1): 3–17.
- Bader, Yousef, and Denise D. Minnis (2000): "Morphological and syntactic code-switching in the speech of an Arabic-English bilingual child". *Multilingua*, 19(4): 383–404.
- Bassiouny, Reem (2006): *Functions of Code Switching in Egypt: Evidence from Monologues*. Boston: Brill.
- Bentahila, Abdelali, and Eirlys E. Davies (1992): "Codeswitching and language dominance". In R. J. Harris, ed., *Cognitive Processing in Bilinguals*. North-Holland Distributors for the U.S. and Canada: Elsevier Science Publishing, 443–457.
- Bentahila, Abdelali, and Eirlys E. Davies. (1991): *Constraints on Codeswitching: A Look beyond Grammar. Vol. 2*. Barcelona: European Science Foundation.
- Bentahila, Abdelali, and Eirlys E. Davies. (1983): "The syntax of Arabic-French codeswitching". *Lingua*, 59(4): 301–330.
- Boumans, Louis (1998): *The Syntax of Codeswitching: Analysing Moroccan Arabic/Dutch Conversations*. PhD dissertation. Tilburg University Press.
- Boussofara-Omar, Naima (2003): "Revisiting Arabic diglossic switching in light of the MLF model and its sub-models: The 4-M model and the abstract level model". *Bilingualism: Language and Cognition*, 6(1): 33–46.
- Budzhak-Jones, Svitlana (1998): "Developing diagnostics: Word-internal code-switching versus borrowing". *Cahiers Linguistiques d'Ottawa*, 26: 1–14.
- Budzhak-Jones, Svitlana, and Shana Poplack (1997): "Two generations, Two strategies: The fate of bare English-origin nouns in Ukrainian". *Journal of Sociolinguistics*, 2(1): 225–258.
- Eid, Mushira (1992): "Directionality in Arabic-English codeswitching". In A. Rouchdy, ed., *The Arabic Language in America*. Detroit: Wayne State University Press, 50–71.
- Eze, Ejike (1998): "Lending credence to a borrowing analysis: Lone English-origin incorporations in Igbo discourse". *International Journal of Bilingualism*, 2(2): 83–201.
- Ferguson, Charles A. (1959): "Diglossia". *Word*, 15: 325–340.
- Fishman, Joshua. (1967): "Bilingualism with and without diglossia; diglossia with and without bilingualism.". *Journal of Social Issues*, 23(2): 29–38.
- Galegher, Erica Lynn. (2012): *Thank You, Merci, Shukran! Private Education and Language Egypt*. Cairo: The American University in Cairo. <https://dar.aucegypt.edu/handle/10526/3135>.
- Haeri, Niloofer (1997): "The reproduction of symbolic capital: Language, state, and class in Egypt". *Current Anthropology*, 38(5): 795–816.
- Hary, Benjamin. (1996): "The Importance of the Language continuum in Arabic multiglossia". In A. Elgibali, ed., *Understanding Arabic: Essays in Contemporary Arabic Linguistics in Honor of El-Said Badawi*. Cairo: The American University in Cairo Press, 69–90.
- Heath, Jeffrey (1989): *From Code-Switching to Borrowing: Foreign and Diglossic Mixing in Moroccan Arabic*. London: Kegan Paul International.
- Hussein, Riyad Fayez, and Ghazi A. Shorrab (1993): "Syntactic constraints on the code switching of Arabic-English bilinguals". *International Review of Applied Linguistics in Language Teaching*, 31(3): 236–241.
- Jake, Janice (1994): "Intrasentential codeswitching and pronouns: On the categorial status of function elements". *Linguistics*, 32: 271–298.
- Jake, Janice L. and Carol Myers-Scotton (2009): "Which Language? Participation potentials across lexical categories in codeswitching". In L. Isurin et al., eds., *Multidisciplinary*

- Approaches to Code Switching*. Amsterdam/Philadelphia: John Benjamins Publishing, 207-242.
- Koning, Anouk de (2009): *Global Dreams: Class, Gender, and Public Space in Cosmopolitan Cairo*. Cairo: The American University in Cairo Press.
- Lewko, Alexander (2012): *Linguistic Projection and the Ownership of English: Solidarity and Power with the English Language in Egypt*. MA dissertation. Cairo: The American University in Cairo Press.
- Mehrez, Samia. (2010): *Egypt's Culture Wars: Politics and Practice*. Cairo: The American University in Cairo Press.
- Mejdell, Gunvor (2006): *Mixed Styles in Spoken Arabic in Egypt: Somewhere between Order and Chaos*. Boston: Brill.
- Mitchell, Terence F. (1986): "What is educated spoken Arabic?" *International Journal of the Sociology of Language*, 16(1): 7-32.
- Mohamed, Fatma. (1989): *Arabic-English Code-Switching in the Speech of an Egyptian Child*. University of Pennsylvania.
- Mustafa, Zahra, and Mahmoud Al-Khatib (1994): "Code-mixing of Arabic and English in teaching science". *World Englishes*, 13(2): 215-224.
- Mustafawi, Eiman (2002): "Lone English-origin nouns in Arabic: Codeswitchers or borrowings?" In S. Burelle and S. Somesfalean, eds., *Actes Du Congrès Annuel de l'Association Canadienne de Linguistique 2002/Proceedings of the 2002 Annual Conference of the Canadian Linguistic Association*. Université du Québec à Montréal, Département de linguistique et de didactique des langues, 219-231.
- Muysken, Pieter (2000): *Bilingual Speech: A Typology of Code-Mixing*. Kindle Edition, Cambridge: Cambridge University Press.
- Muysken, Pieter (1997): "Code-switching processes: alternation, insertion, congruent lexicalization". In M. Pütz, ed., *Language Choices: Conditions, Constraints, and Consequences*. Amsterdam/Philadelphia: John Benjamins Publishing, 361-380.
- Muysken, Pieter (2007): "Mixed codes". In P. Auer and W. Li, *Handbook of Multilingualism and Multilingual Communication*. Berlin: Mouton de Gruyter, 315-339.
- Myers-Scotton, Carol (1998): "A way to dusty death: the matrix language turnover hypothesis". In L. A. Grenoble and L. Whaley, eds., *Endangered Languages*. Cambridge: Cambridge University Press, 289-316.
- Myers-Scotton, Carol (1992): "Comparing codeswitching and borrowing". In C. M. Eastman, ed., *Codeswitching*. Clevedon, Philadelphia, Adelaide: Multilingual Matters Ltd, 19-39.
- Myers-Scotton, Carol (1993): *Duelling Languages: Grammatical Structure in Codeswitching*. Oxford: Oxford University Press.
- Myers-Scotton, Carol. (2006): *Multiple Voices: An Introduction to Bilingualism*. Oxford: Blackwell Publishing.
- Myers-Scotton, Carol (2010): "Patterns and predictions for code-switching with Arabic". In R. Bassiouney, ed., *Arabic and the Media*. Leiden: Brill, 81-95.
- Myers-Scotton, Carol, and Janice Jake (2009): "A Universal model of code-switching and bilingual language processing and production". In B. E. Bullock and A. J. Toribio, eds., *The Cambridge Handbook of Linguistic Code-Switching*. Cambridge: Cambridge University Press, 336-357.
- Myers-Scotton, Carol et al. (1996): "Arabic and constraints on codeswitching". In M. Eid and D. Parkinson, eds., *Perspectives on Arabic Linguistics IX: Papers from The Ninth Annual Symposium on Arabic Linguistics*. Amsterdam/Philadelphia: John Benjamins Publishing, 9-43.

- Nortier, Jacomien (1990): *Dutch-Moroccan Arabic Code-Switching among Young Moroccans in the Netherlands*. Dordrecht: Foris.
- Okasha, Maha (1999): *Structural Constraints on Arabic/English Codeswitching*. PhD dissertation. South Carolina: University of South Carolina.
- Poplack, Shana (1980): "Sometimes I'll start a sentence in Spanish Y TERMINO EN ESPAÑOL: Toward a Typology of Code-switching". *Linguistics*, 18(7/8): 581-618.
- Poplack, Shana (1993): "Variation theory and language contact". In D. R. Preston, ed., *American Dialect Research*. Centennial Series of the American Dialect Society, Amsterdam/Philadelphia: John Benjamins Publishing, 251-286.
- Poplack, Shana, and Marjory Meechan (1998): "Introduction: How languages fit together in codemixing". *International Journal of Bilingualism*, 2(2): 127-138.
- Poplack, Shana et al. (1988): "The social correlates and linguistic processes of lexical borrowing and assimilation". *Linguistics*, 26: 47-104.
- Poplack, Shana et al. (1989): "Distinguishing language contact phenomena: evidence from Finnish-English Bilingualism". *World Englishes*, 8(3): 389-406.
- Rouchdy, Aleya (1992): "Borrowing in Arab-American Speech". In A. Rouchdy, ed., *The Arabic Language in America*. Detroit: Wayne State University Press, 36-9.
- Sallo, Ibrahim Khidhir (1994): "Arabic-English code-switching at the University: A Sociolinguistic Study". In R. de Beaugrande et al., eds., *Language, Discourse and Translation in the West and Middle East*. Amsterdam/Philadelphia: John Benjamins, 115-131.
- Sankoff, David, et al. (1990): "The Case of the nonce loan in Tamil". *Language Variation and Change*, 2: 71-101.
- Schaub, Mark (2000): "English in the Arab Republic of Egypt". *World Englishes*, 19(2): 225-238.
- Turpin, Danielle (1998): "Le Français, c'est le last frontier": The Status of English-Origin Nouns in Acadian French". *International Journal of Bilingualism*, 2(2): 203-219.

How *VIPs* and *IT-professionals* behave in Russian. The use of loan abbreviations in Russian analytical composites

Elizaveta Tarasova
IPU New Zealand Tertiary Institute
etarasova@ipu.ac.nz

ABSTRACT

The current research is focused on the English N+N pattern, which is widely reported to have been borrowed into Russian under the influence of numerous compounds that have entered Russian recently. The data used for analysis include loanwords with N+N structure (e.g., *бизнесмен* [businessman], *арт-магазин* [art shop]), a considerable proportion of which are units with loan abbreviations in the leftmost (modifying) position (e.g., *пиар-директор* [pi'ar direktor] 'PR director', *ВИП-места* [vip-mesta] 'VIP seats'). The paper looks at the general principles of assimilation using Krysin's (1975) framework to understand whether it is possible to talk about changes in the Russian morphology under the influence of lexical borrowing. The analysis of the Russian N+N data to date allows for two explanations for the growing productivity of the analytical pattern. One of these is the revival of the superficially similar Russian N+N morphological pattern which had fallen out of productivity. The other explanation concerns the necessity of preserving the meaning of the loan and its conceptual value, which may otherwise be affected in the process of derivational assimilation.

Keywords: assimilation, N+N compounds, Russian N+N composites, derivation, loanwords, loan abbreviations

1. Introduction

A continuous interest in the research devoted to the issue of changes in European and world languages under the influence of English has been growing. The topics of investigation range from the patterns of assimilation of loanwords and changes in the



lexis (Haspelmath and Tadmor, 2009; Haugen, 1950; Hyman, 1970) to the changes in morphological processes (Panov, 1999; Ryazanova-Clarke and Wade, 1999; Dimitrova et al, 2011; Gazda, 2008; Rochtchina, 2012; Panov, 1999) and even syntactic structures (Nagano and Shimada, 2016). One of the issues in this area is the tendency to analytism in word-formation processes, which is attributed to the influence of (American) English through the number of borrowings registered in inflectionally rich languages.

The overarching question that this investigation attempts to contribute to is compound genesis in Russian, and this paper aims to present some insights on how compounds that use the English N+N pattern assimilate and/or develop in Russian. Analytical N+N compounds like *бизнес-план* [biznes-plan] ‘business plan’ and *рок-музыка* [rok-muzyka] ‘rock music’ are viewed as items of lexical borrowing, and they are often reported to impact Russian derivational morphology, shifting it towards analytism (Dyakov, 2001; Rochtchina, 2012; Ryazanova-Clarke and Wade, 1999, to name just a few). Similar claims are made about other Slavic languages, e.g. Czech (Gazda, 2008), Bulgarian (Bagasheva, 2016), Polish (Sztencel, 2009), and others. This research looks at the matters of assimilation of English N+N compounds and the extension of the analytical N+N pattern in such an inflectionally rich language as Russian. For the reasons of space, the paper focuses mainly on Russian N+N compounds in which the first element is represented by a loan abbreviation (LA) that functions as a word.

This investigation is part of a corpus-based project in which a database of 2,000 Russian analytical composites was collected and analysed. The types of Russian composites were picked from the *Russian National Corpus* (henceforth – the RNC), as well as from the current Russian printed media; e.g., *Известия*, *Экономист*, *Аргументы и факты*, etc. The choice of the resources for data collection was determined by the necessity of looking at the occurrences of Russian composites in non-specific contexts. This research looks at the data from a synchronic perspective for two reasons. Firstly, the frequency-based evidence is not as important for this research as the analysis of various types of LA+N compounds; and secondly, the RNC is relatively small and cannot be used as a reference point at this stage. However, looking at the matters diachronically would provide a deeper understanding of the development of this morphological pattern.

The analysis of loans and stages of their assimilation in the language uses the criteria offered by Krysin (1975) to decide whether a loan is fully or partially assimilated in the target language. The reason for choosing Krysin’s set of criteria over others—e.g., Haugen’s (1950)¹—is that this set answers the requirements of this study because it was developed for the analysis of loans in Russian and covers the points (shown below) relevant to this investigation. Another reason for choosing this framework for the analysis is dictated by the aim of the paper, which is to look at the overall trend in the morphology of modern Russian rather than assign the analysed compound types to certain categories. Besides, as justly noted by Gómez Capuz (1997: 82), most of the proposed typologies and classifications were created to match mainly Western languages. As a result, they may lack full validity since the discrete categories they operate with cannot account for all the factors involved in a given socio-linguistic situation. Finally, Krysin’s criteria

allow for focusing on the process of assimilation of one and the same borrowed word on different levels (phonological, orthographic, semantic, grammatical).

The paper starts with the general description of characteristics of N+N compounds in English and Russian, and demonstrates the similarities and differences between the patterns of formation to explain how the English N+N pattern is utilised as a special way of assimilating novel loans in general, with special attention to assimilation of LAs.

2. English N+N compounds and compounding in Russian

N+N compounds in English represent an interesting field of research. The issue of the position of N+N compounds in English grammar, as well as other topics, including classification of compounds, their formal structure and semantic characteristics, are rather controversial. For the purposes of this paper, which is not focused on the questions associated with N+N compounding in English, the generic definition that a compound is a new lexeme formed by adjoining two or more lexemes (Bauer, 2003: 40) is accepted. The term “N+N compound” is used to refer to English noun-noun sequences that satisfy the above definition.

Overall, the process of combining two nouns in order to create a new concept is believed to be typical of analytical languages. However, the results of the study by Štekauer et al. (2008) suggest that compounding is widely spread across languages, even though the means for creating nominal compounds may differ. Hence, it is interesting to see how similar/different Russian compounds are from English N+Ns.

2.1. Russian N+N compounds

Shvedova (1980: 139) refers to the word-formation pattern of combinations with the N+N structure as ‘agglutination’/‘pure compounding’, that is a type of word formation where the base component is equivalent to a full word and the preceding component is represented by a root morpheme. The meaning of the new word is described as the sum of the meanings of the components. The first component is believed to be a root morpheme (rather than a word), and the whole construction is a complex word.²

Pure compounding is one of the productive word-formation processes in Russian, in which two noun roots/bases are normally connected (“glued” to each other) by an interfix *-o-*, as in *хлебозавод* [khlebozavod] ‘bread-making plant’ or *-e-*, as in *птицефабрика* [ptitsefabrika] ‘poultry factory’. Compounds that employ zero interfix are usually characterised by the neuter gender of the first element; e.g., *радиостанция* [radiostantsiya] ‘radio station’. Neuter gender nouns in Russian end in either *-e* or *-o*; therefore, it is not always clear whether the final vowel of such nouns belongs to the first element of the compound or whether this vowel is an interfix. Units formed by the process of pure compounding also include cases which have formal characteristics of English N+Ns, whereby the first element is represented by a free root morpheme that can be used as a noun lexeme (N1). The use of the hyphen or one-word spelling in such sequences is

often considered as the main criterion for identifying whether we deal with a complex word rather than a phrase (Shvedova, 1980).³ However, this type is unproductive and the formations are limited to very few Russian words which can be used in N1 position, e.g. *царь*, as in *царь-пушка* [tsar'-pushka] 'king cannon'; *жар*, as in *жар-птица* [zhar-ptitsa] 'fire bird'⁴; *горе*, as in *горе-человек* [gore-chelovek] 'unfortunate person'; and *чудо*, as in *чудо-печь* [chudo-pech'] 'wonder oven' (Shvedova, 1980).

The important characteristic of such sequences which allows for drawing a parallel between Russian complexes and English N+N compounds is the lack of inflectional markers on the leftmost element (e.g. for case and number) when it is required by syntax. This means that N1 in such units is used analytically, which makes them structurally similar to English N+Ns.

Когда после сорокалетнего перерыва открыли Кремль для всеобщего обозрения, я, ..., в месте со своими друзьями кинулся скорей посмотреть *царь-пушку* ([tsar'-pushk.u] 'tsar'-cannon:SuffAcc') и *царь-колокол*.

(When, after a forty-year break, the Kremlin was opened for public viewing, I, ..., together with my friends rushed to see the **tsar cannon** and the tsar bell.)

Я давно собирался посмотреть на гнездышко такой *жар-птицы* ([zhar-ptits.y] 'fire-bird:SuffGen'), как Вы, — шутя сказал Иван Павлович.

(I've been going to look at the nest of a **firebird** like you for a while - Ivan Pavlovich said jokingly.)

The examples above demonstrate the analytical use of the first component in the Russian N+N compounds, where the leftmost component has a zero ending, and the obligatory case marking is assigned to the head. Hence, it is possible to suggest that one of the compound formation processes in Russian employs the pattern which bears formal similarity to the analytical English N+N compounds.

The limited number of analytical N+N compounds in Russian can be explained by the fact that being a highly inflected language, Russian allows for the formation of adjectives from practically any noun, verb or other adjective (Townsend, 1968; Corbett, 2004). Consequently, indigenous Russian words and well-assimilated loans may serve as derivational bases for the formation of adjectives. As Corbett (2004: 201) notes, Russian employs a derived adjective + noun sequence *книжный магазин* [knizh.nyi magazin] 'bookshop', while many European languages use an N+N compound. Therefore, there is no pragmatic need for the use of nouns for modification, and availability of adjectives makes modifying a noun using another noun redundant.

2.2. Borrowed N+Ns in Russian

In recent years, the number of items that use the English N+N pattern has increased dramatically, and a considerable amount of research on this issue has started to appear. The term that is often used for these constructions is "analytical composites" (Karaulov, 1998: 443) because the modifying element in composite structures does not acquire inflections to reflect the syntactic function, part of speech and/or case. The increase in

these units is attributed to the influence of borrowed English N+Ns, in which the first element bears no morphological markers to allow for defining its morphological category. Russian N+N composites use loanwords in the modifier position, but some of them have been created in Russian rather than borrowed from English (e.g. *бизнес-леди* [biznes-ledy] ‘business woman’). This gives rise to the suggestion that borrowing English compounds into Russian may have also triggered borrowing the word-formation pattern. At the same time, it may also be the case that assimilation of the N+N compounds could be viewed as a catalyst for bringing back to life the old *царь-пушка* [tsar’-pushka] ‘king cannon’ pattern, which has fallen out of productivity.

The paper analyses Russian analytical N+N composites (specially focusing on the use of LAs in the modifying position) to see how borrowed units with this structure develop in Russian and whether it is justified to speak about borrowing of the morphological pattern under the influence of lexical borrowing. Before moving to the analysis of the Russian N+Ns, it is important to outline how the principles of assimilation work in Russian.

3. Assimilation of loans in Russian

The analysis of loanwords needs to consider the stages a loan undergoes in the process of assimilation in the new linguistic environment. The chosen framework (Krysin, 1975) addresses the following criteria:

1. Categorising the graphic-phonetic representation of a loan by means of the spelling system of the recipient language (phonological and orthographic assimilation);
2. Referring the borrowed lexeme to a certain grammatical class and category as prescribed by the grammar of the recipient language (grammatical assimilation);
3. Confirming semantic independence of the borrowed word when a loan is not a full synonym of some indigenous word of the recipient language or an earlier borrowing from elsewhere that is now accepted as part of the recipient language (semantic assimilation);
4. Identifying the derivational productivity of the borrowed word (derivational assimilation).

Below is a brief description of different kinds of assimilation of loans in Russian. Derivational assimilation is discussed in more detail for reasons that will be explained later in the paper.

3.1. Phonological assimilation

Once a borrowed word enters a language, it changes its phonological form in accordance with the restrictions on phonemes and their distribution imposed by the phonetic, phonological and phonotactic systems of the recipient language. Foreign sounds are

replaced with similar sounds of the recipient language to keep the loan phonologically close to the original and to make its phonological shape more native-sounding. This phenomenon is referred to as ‘phoneme substitution’ (Broselow, 1992: 200). Haspelmath (2009) gives the example of the loan *резюме* [rez’ume]Fr. ‘résumé’ in Russian. Since Russian does not have front rounded vowels, French [y] turns into [ю], and the preceding consonant becomes palatalised.

3.2. Orthographic assimilation

The differences in the English and Russian writing systems determine the character of the orthographic representations of loans. Spelling of loans depends on the channel of borrowing: oral or written (Volodarskaya, 2002: 106). Speakers of the recipient language may use transcription (*афтешеѝв* [afte shejv] ‘after shave’), transliteration (*фронт-офис* [front-ofis] ‘front office’), or a combination of the two (*менеджер* [menedzher] ‘manager’).

In a limited number of cases (mostly abbreviations), Latin script is used for loans in Russian (*VIP, CD, SMS, PIN, PUC, HR, PR*, etc.), which sometimes function on a par with the transcribed/transliterated variants⁶:

*VIP – ВИП/вип, CD – суду/СуДи, SMS – эсэмэс, PIN – ПИН/пин, PUC – ПУК, HR – эйчар, PR – нуар*⁷.

3.3. Grammatical assimilation

Grammatical assimilation involves assigning grammatical categories to loans. Russian is a language with a highly developed system of flexions that are used to denote grammatical categories. Each new word needs to be referred to one of the grammatical classes irrespective of whether the same categories exist in the donor language and sometimes irrespective of the linguistic means the donor language uses. For example, the loan *байк* [baik] ‘bike’ is referred to 2nd Declension Masculine nouns because it does not have a suffix at the end similar to Russian nouns that belong to this category.

3.4. Semantic assimilation

When a word is borrowed, it is usually the case that only one sense of a possible plethora of meanings is borrowed at a time (Busse and Görlach, 2002: 27). The loan’s integration into a semantic field may involve a change of its meaning, both in denotation and connotation (Ibid.), which is commonly called semantic assimilation. In the process of semantic assimilation, the original meaning of the word may be either extended or specified (Kimyagarova, 1968). The case of the loan *паркинг* [parking] ‘parking’ listed as “a place/building for paid secure parking” (Komlev, 2006)⁸ as opposed to *парковка* [parkovka] ‘parking’, the noun motivated by an earlier loan *парковать* [parkovat] ‘to park’ registered in the meaning “bringing a car to a standstill” (Kotelova, 1984), is a good example of that.

3.5. Derivational assimilation

Derivational assimilation of a loan refers to its involvement in word-formation processes (Volodarskaya, 2002: 109), when it starts being used as a derivational base motivating new words and building different kinds of word families: word chains and word nests (Tikhonov, 2014: 621-622). Once loans start being involved in derivation, they participate in the sense creation process in a similar way as other words of the language. Concepts that are biologically, socially or culturally meaningful for a society are employed in the word-formation processes, i.e. are “derivationally marked” (Vendina, 1999: 16). Native speakers of the recipient language create derivatives using loans as bases according to their understanding of how the system of the language works. They then give new names to the objects that are important for them using this knowledge.

The choice of derivational material used for deriving new words from loans is believed to be important. Appearance of loan-based derivatives that carry connotational meaning not registered in a foreign word demonstrates the values of the conceptual content of a loan which native speakers of the recipient language accentuate by means of derivation. This can be noticed in hybrid formations, in which foreign roots are combined with indigenous Russian affixes (especially those primarily used in colloquial speech). As stated by Townsend (1968), the overall tendency in Russian word-formation is that foreign derivational means are usually employed for forming derivatives from foreign roots. For example, the blended derivative *программист* [programmist] ‘computer programmer’ motivated by the loanword *программа* [programma] ‘programme’ uses the foreign suffix *-ист* [-ist] and is stylistically neutral. Deviation from this tendency and the use of a Russian suffix *-щик* [-shchik], e.g. *програмщик* [programshchik] ‘programmer’ (colloquial), results in the change of the stylistic status of the concept the derivative denotes and produces a humorous effect, thus downgrading the value of the concept.

An interesting case for analysis is presented by the derivatives formed on the basis of the compound *бизнесмен* [biznesmen] ‘business man’. Even though it seems derivationally analysable due to the recurrence of the first element in a number of *бизнес+N* ‘business+N’ compounds in Russian, it was borrowed as a single unit and is assimilated accordingly. Almost all derivatives based on this loan belong to expressive vocabulary and convey negative connotations: *бизнесменишка* [biznesmen-ishk-a] ‘small, unsuccessful businessman’, rude; *бизнесменчик* [biznesmen-chik] ‘a businessman with a very small business’, humorous; *бизнесмениша* [biznesmen-sha] ‘business woman’, rude; *бизнесменка* [biznesmen-ka] ‘business woman’, humorous; *бизнесменствовать* [biznesmen-stvo-vat] ‘do business’, humorous; *бизнесменствующий* [biznesmen-stvu-yushch-yi] ‘doing business’, humorous. The humorous effect in the first two examples (*бизнесменишка* and *бизнесменчик*) can be explained by the use of diminutive suffixes. In the case with *бизнесмениша* and *бизнесменка* the presence of additional connotations can be attributed to the use of gender-marking suffixes, implying that business is not suitable for a woman. However, the last two cases *бизнесменствовать* and *бизнесменствующий* employ stylistically neutral suffixes.

This allows for a speculation that the use of indigenous Russian suffixes may be responsible for the derivatives of most recent borrowings becoming stylistically marked even in cases when the suffix is productive and forms neutral derivatives in Russian. For example, the agentive suffixes *-щик* [-shchik] and *-ник* [-nik] are used to form nouns from verbs to name people or things related to activities named by the root, e.g. *танцовщик* [tantsovshchik] ‘dancer’, *буксировщик* [buksirovshchik] ‘towing machine’, *шутник* [shutnik] ‘joker’, *работник* [rabotnik] ‘worker’. When used in combination with a recently borrowed root, the derivatives may sound odd, especially in cases when a more acceptable form with a foreign suffix already exists in the language, as is the case with *программист* [programmist] ‘computer programmer’. Adding suffixes to mark loans for feminine gender produces a similar effect, e.g. *блоггерша* [bloggersha] ‘female blogger’, *интернетчица* [internetchitsa] ‘female internet user’.

Even though the research that looks at hybridisation as the means of assimilation of loans seems to be on the rise in Russian linguistics, there does not seem to be much work devoted to the influence of the use of different derivational means on the assimilation of loans. A study that would address this issue would be beneficial for better understanding of peculiarities of assimilation mechanisms in Russian.

4. English N+N compounds in Russian

We have seen that assimilation of a loan in a language undergoes stages, with derivational assimilation usually being the last. It is difficult to predict how long it may take for a word to be fully assimilated in a language, since several factors are involved, including the formal characteristics of the word, its communicative value, frequency of occurrence, the register in which the word is used, and the importance of the referent for the speakers. If we assume that N+N compounds are words, the question we can ask is: what does it take for such words to become assimilated? Based on the discussion above, one of the most important markers of successful assimilation of a loan is derivational assimilation. We have seen, however, that Russian has extensive derivational means for individual words to be successfully assimilated in the language, so the reasons for borrowing N+N compounds and using the compound formation pattern are somewhat unclear. The discussion below looks at the possible explanations for this phenomenon.

4.1. Russian N+N data

As noted in the Introduction, this investigation is part of a larger corpus-based project in which a database of 2,000 Russian analytical composites was collected and analysed. What follows is a brief outline of the whole study, then the discussion is narrowed down to the analytical N+N composites with a loan abbreviation in the modifying position.

The criteria for the Russian N+N composites to be included in the analysis are as follows:

1. The first element of the analytical composite is a loanword, which was borrowed and/or came into productive use relatively recently (starting from the 1990s); e.g.,

интернет[internet] ‘the Internet’, *фешн* [feshnn] ‘fashion’. Earlier loans like *бизнес* [biznes] ‘business’ and *риск* [risk] were included if they were borrowed as part of N+N compounds to refer to novel concepts, e.g. *бизнес-решение* [biznes-resheniye] ‘business solution’, *риск-менеджмент* [risk-menedzhment] ‘risk-management’.

2. The first element must be used analytically (i.e. without grammatical markers) in the composite irrespective of the syntactic position, gender of the noun and plurality of the head.
3. Neither element of the composite should be unique; i.e., elements recur in the language in the same structural form and the same (or similar) meaning either as free lexemes or as components of other composites.
4. The composite should be semantically and morphologically transparent. Examples like *пинг-понг* [ping-pong] ‘ping-pong’, *кьюр-цензура* [kuryur-tsenzura] ‘censorship’, *тай-брейк* [tai-breik] ‘tiebreak’ were borrowed into Russian as chunks to denote the concepts that were new for Russian culture. They appear to be morphologically divisible; however, this division is only possible because of the orthographic representation.
5. The analytical element cannot be a truncated adjective/noun. The reason for including this criterion into the list is determined by the fact that the *truncated adjective + noun* pattern is very productive in Russian. This criterion was the most difficult to observe since in some cases (e.g., *спорт+N*(sport+N) composites) defining the first composite as a noun (and not a truncated form from the adjective *спортивный* [sportivnyi] ‘relating to sport’) can only be based on the etymological data (Chernykh, 1999; Fasmer, 1986), which is not always complete and/or precise.

Observation of these criteria was important for the research since they allow for deeper understanding of the reasons for N+N compounds (rather than individual words) to be borrowed into Russian.

Several constituent families that satisfy the search criteria and that have the largest number of composite types were chosen for the analysis. The leftmost elements of the collected composites were approached from the viewpoint of their morphology to determine the factors that encourage the formation of analytical constructions. It was noticed that most N+N analytical composites contain loanwords both as the head and as the modifier. It was pointed out above that once loanwords start motivating the formation of new words in the target language, the derivational material (suffixes and prefixes) used for building paradigms with foreign roots are usually of foreign origin (Townsend, 1968). The same seems to hold true for the formation of composites containing a loan element, and the number of novel coinages with both words being loans should then be expected to be higher than the number of formations in which one element is a loanword and the other one is a Russian word.

The head element in the collected sample is usually represented by a grammatically and derivationally assimilated loan which often has lost its novelty in the language; e.g.,

интернет-технология [internet-tekhnologiya] ‘internet technology’, *видеосалон* [videosalon] ‘video salon’, *онлайн-конференция* [onlain-konferentsiya] ‘online conference’.

The leftmost (analytical) element in more than 90% of the composites under discussion is represented by the latest loanwords, borrowed mainly from English. The analytical elements usually refer to notions that have become important in recent times. With the disappearance of the “iron curtain”, Russian culture was flooded with objects, concepts and phenomena that were not accessible to people during Soviet rule. The massive number of loans from English at the end of the 20th century can be explained by the radical changes in the life of the country. Since Russia was isolated from the Western world for most of the century, the information that was coming from behind the “curtain” was strictly censored and the Western lifestyle was severely criticised. Therefore, developments in many spheres of life, changes in attitudes to political, societal and economic matters, and achievements in culture and science were new to the Russian lifestyle once the political and social situations in the country changed.

According to the collected corpus, the spheres of entertainment and business have acquired the largest number of loans. The other areas include: sports (*фитнесс-инструктор* [fitness-instruktor] ‘fitness instructor’, *допинг-контроль* [doping-kontrol’] ‘doping control’); fashion (*имидж-студия* [imidzh-studiya] ‘image studio’, *фешн-шоу* [feshn-shou] ‘fashion show’); and IT (*веб-сайт* [veb-sait] ‘website’, *интернет-конференция* [internet-konferentsiya] ‘online conference’). Since the examples for this research were collected from current Russian print media (newspapers and magazines) aimed at a general audience, the data cannot accurately reflect changes in more specific areas.

The collected examples fall into three distinct groups:

- (1) The first element in the composites has the characteristics of a 2nd Declension, masculine noun. This group is the biggest in terms of the number of examples and in terms of the variety in the loan components. The composites of this group are structurally similar to Russian compounds with a first element truncated, e.g. *спорт-площадка* [sport-ploshchadka] ‘sports grounds’ from *спортивная площадка* [sportivnaya ploshchadka] ‘sports grounds’.
- (2) The first element in the composites ends in a vowel *-o*, *-e* or *-a*, and is morphologically unassimilated (i.e., it often does not have an inflectional paradigm). The composites of this group are structurally similar to Russian compounds with a connecting vowel, e.g. *хлебозавод* [khlebozavod] ‘bread-making plant’.
- (3) The first element is represented by a loan abbreviation (henceforth – LA) that acts as a word, e.g. *вип-зона* [vip-zona] ‘VIP zone’.

This investigation focuses mainly on the third group and attempts to analyse the assimilation patterns that LAs demonstrate. The main reason for choosing this group is that LA+N composites are the most recent phenomenon and thus provide a clearer picture

for analysis. The next part of this paper looks at how the system of the Russian language accommodates LAs through their use in analytical composites.

4.2. Abbreviation

“Abbreviation” (or shortening), in the most general sense, is a type of word-formation in which the elements are reduced to the initial letters and/or syllables of constituting lexemes (Crystal, 2008). Abbreviation serves as an example of the tendency to an economical use of language, thus raising the communicative value of language items. Abbreviation is an umbrella term for formations that include initialisms (*USA, FBI*), acronyms (*NATO, UNO*), and clippings (*Fro* from *Afro*, *op art* for *optical art*).⁹ Initialisms are opposed to acronyms, which are abbreviations pronounced according to the rules of the phonological system of the language, e.g. *NATO* [ˈneɪtəʊ] (for North Atlantic Treaty Organisation), *ANZAC* [ˈænzæk] (for Australian and New Zealand Army Corps), *OPEC* [ˈəʊpɛk] (for Oil Producing and Exporting Countries). Both acronyms and initialisms are products of word-formation, since they possess lexical meaning and can be employed for further derivational operations (in the same way as coinages and loans). At the initial stages of their formation, the meaning of abbreviations is motivated by the meaning of the words that constitute them; this motivation tends to be lost with time. The term “abbreviation” will be used to refer to two main types that are analysed in the current research—that is acronyms and initialisms—because these are the two main types found in the collected data set.

4.2.1. Lexicalisation of abbreviations

If we assume that abbreviations function in the language in a similar way to regular words, it is possible to speak about their lexicalisation. In the case of lexicalised abbreviations, there is no need for analysis of the compressed structure; we are taken straight to the concept that is labelled with an abbreviated form. The more frequently an abbreviation is used, the sooner its referent becomes known to the language users, and the sooner it loses its connection with the motivational base. The process of demotivation of abbreviations may be followed by their orthographic and derivational assimilation. In this case abbreviated lexemes may start being used as derivational bases and start forming paradigms. The word *laser* (*Light Amplification by Stimulated Emission of Radiation*), demonstrates how the term may lose its connection with the motivational base, because the latter is not important for native users of the language. It has the spelling of a regular word, and it is used in two meanings: ‘a kind of beam’ and ‘a unit that emits this type of beam’ (neither of which has the reference to the process, originally present in the motivational base). It also has forms *laser* and *lasers* and is perceived as a cognitively whole notion.

As English is an analytical language, the tendency for lexicalised abbreviations to be used as motivational bases is not very strong even in the case of acronyms. The derivational activity of these lexical units is revealed in the formation of compounds in

English where they realise the derivational potential by forming extended paradigms; e.g., *US Army, US Navy, US government, US Cabinet; FBI certificate, FBI agent, FBI recruitment policy, FBI unit*.

4.3. LA+N composites

The current investigation looks at the results obtained from the analysis of around 600 types of Russian analytical composites, in which the first element is represented by a loan abbreviation (LA). The purpose of the analysis is to look at the ways in which LAs are assimilated, with special attention to their use in analytical composites.

The first question to ask is whether LAs have the potential to be assimilated in the recipient language (Russian) and what it means for an LA to become assimilated. When an abbreviation is borrowed, it is perceived by the speakers of the recipient language as an indivisible lexical unit, which requires an explanation since the concept it refers to may not be familiar to the speakers.

В Америке их около десятка — кроме упомянутых ЦРУ и ФБР, ещё **АНБ (Агентство национальной безопасности)**, **РУМО (Разведывательное управление Пентагона)**, действующие автономно разведки различных видов вооружённых сил, информационная служба государственного департамента и немало других ведомств.

There are about a dozen of those in America, apart from those mentioned; i.e., the CIA and the FBI, there is also **the NSA (National Security Agency)**, **the DIA (Defence Intelligence Agency)**, some autonomously acting intelligence agencies of different kinds of armed forces, the information service of the Department of State, as well as many other departments.

The examples above¹⁰ may be used to demonstrate how the well-known abbreviations referring to CIA and FBI are opposed to NSA (National Security Agency) and DIA (Defence Intelligence Agency), which require an explanation. If the abbreviation is used frequently and its referent is known to the majority of speakers (which is often determined by the longevity of the LA in the language and its frequency of occurrence in general contexts), the explanation is not required. This process is the same for loanwords, especially in the case of “loanwords by necessity”, or “cultural borrowings (Haspelmath, 2009: 46).

Another argument supporting the suggestion that abbreviations (and LAs) are perceived as indivisible units is their use in the so-called tautological compounds; e.g., *VIP-персона* [VIP-persona] ‘VIP-person’, *SMS-сообщение* [SMS-soobshchenie] ‘SMS-message’, *IT-технология* [IT-tekhnologiya] ‘IT-technology’.¹¹ In such compounds, an abbreviation loses its connection to the motivational base (words comprising the abbreviation), which justifies the fact that abbreviations tend to become lexicalised and be perceived as single lexical units.

4.4. Assimilation of loan abbreviations

If LAs are perceived as words, it is possible to suggest that they function as words too. Therefore, the process of their assimilation should match the assimilation of regular loans. The same criteria should be in place for determining the potential for an abbreviation to be assimilated in the recipient language. An assimilated abbreviation should thus:

- Be orthographically represented by the means of the recipient language (*ФБР* for FBI).
- Be assigned a grammatical category (*ФБР* is tagged as an inanimate, singular noun of neuter gender).¹²
- Be semantically assimilated. Since the loan *ФБР* denotes a foreign organisation, we cannot see the semantic extension. However, no explanation of meaning for *ФБР* is given in the analysed texts, which suggests that it is semantically assimilated.
- Participate in the derivational processes.

The status of the loans that are abbreviations in the source language is not always clear in Russian. Based on their degree of assimilation, it is possible to identify the following groups.

4.4.1. Transplants

LAs that use the orthography of the donor language are often called transplants or barbarisms. The tendency for LAs to be used in their original spelling is quite strong and is preserved even when forming composites, e.g. *PC-система* [PC-systema] ‘PC system’, *VIP-зал* [VIP-zal] ‘VIP hall’, *CD-плеер* [CD-pleyer] ‘CD player’, *PR-кампания* [PR-kampaniya] ‘PR campaign’, etc.

4.4.2. Graphically assimilated abbreviations

LAs that acquired Russian orthographic representation are graphically assimilated. Some of them demonstrate the principles of transcription, preserving the pronunciation of the etymon, e.g. *сиду* (from CD), *нису* (from PC), *нуар* (from PR). Others demonstrate the principles of transliteration, e.g. *вун* [vip] (from VIP). It is also often the case that one and the same concept can be represented in the language by a graphically assimilated abbreviation and a transplant. For example, the LA *вун* can occur both in Russian and English orthography.¹³

4.4.3. Lexically assimilated abbreviations

The next stage of assimilation of LAs refers to the condition in which the constituents comprising a loan are translated; e.g., *ЦПВ* (for CIA), *ФБР* (for FBI), *ПК* (for PC). In Haugen’s (1950) classification such loans are viewed as kinds of semantic borrowing, which includes loan translation (calques), loan rendition and loan creation (Fischer, 2008: 6). In the case with lexically assimilated abbreviations, it is possible to suggest that we

deal with calques; however, the syntactic nature of the source phrases is not homogeneous.¹⁴ In order to avoid confusion, a more generic term is used here.

4.4.4. Grammatically assimilated abbreviations

When we talk about abbreviations, the distinction between grammatical assimilation and derivational assimilation is hard to draw in Russian. Once an LA starts developing an inflectional paradigm, it is possible to suggest that it becomes grammatically assimilated; e.g., *випы* [vipy] ‘VIP.PIGen’, *СПИДа* [SPIDa] ‘AIDS.Gen’. Assimilation on the grammatical level does not always require assimilation on the lexical level (translation of the constituents). For example, such LAs as *ВИП/вип* [VIP] ‘VIP’ and *пиар*[piar] ‘PR’ are not lexically assimilated, but they may still be marked for case and gender when required by syntax of the construction.¹⁵

У нас просто меньше лицемерия и **пиара**. (*piar.a* [piar.Gen])
(We just have less of hypocrisy and PR.)

Мероприятие представляет собой танцевальную промо-акцию для журналистов, **ВИПов**(*VIP.ov* [VIP.PIGen]), представителей прессы, промоутеров, диджеев...

(The event represents in itself a dance promotion campaign for journalists, VIPs, media representatives, promoters, DJs...)

Overall, the number of grammatically assimilated LAs is quite low in the collected data set. Examples include *BASIC*, *PIN*, *VIP*, and some others.

4.4.5. Derivationally assimilated abbreviations

Once an abbreviation starts being used as a derivational base, we can talk about its derivational assimilation. Derivatives that are formed on the basis of LAs may include units at any stage of assimilation starting from transplants: *ВИПовский* [VIPovskiy] ‘typical of/aimed at VIPs’ and orthographically assimilated units: *пиарщик* [piarshchik] ‘PR-specialist’ to lexically assimilated units: *ФБРовцы*[FBRovtsy] ‘FBI officers’. The number of abbreviations that are used as derivational bases is limited, which is probably determined by the generally low derivational potential of abbreviations. LAs are widely used in LA+N sequences instead. The reasons for this and some implications are discussed below.

4.5. Peculiarities of assimilation of loan abbreviations

As shown above, assimilation of a loanword follows distinct stages. The overall tendency is that one stage includes the previous; i.e., from phonological assimilation to derivational assimilation. This means that a derivationally assimilated loan needs to be assimilated on other levels as well. This does not seem to be the case for abbreviations. The analysis of abbreviation-based derivatives shows that the process is somewhat sporadic. For

example, it is noticed that a transplanted form and its russified word-like version function in the language simultaneously as the examples from the RNC (n.d.) demonstrate.

PRvs. пиар (piar)

Только на наружную рекламу, социологические исследования и **PR** было потрачено не менее 8 млн.

(No less than 8 million was spent just on outdoor advertising, case studies and PR.)

Что сегодня первично: война или **пиар** — понять довольно сложно.

(It is hard to understand what comes first nowadays: war or PR)

вип/ВИП vs. VIP

И если не решить этот вопрос кардинально, очередная **вип**-авария может иметь самые неприятные политические последствия.

(And if we do not resolve this issue, another VIP-accident may have most unpleasant political consequences.)

Кроме высоких гостей президента Путина, в город зачастили и другие **VIP**-персоны.

(In addition to President Putin's high guests, other VIPs started to visit the city quite often.)

SMS/smsvs. СМС

А в Москве в ближайшее время впервые в стране начнет работать служба **sms**-оповещения о состоянии воздуха в разных районах города.

(In the near future in Moscow an sms-alert service informing about air quality in different areas of the city will start functioning for the first time in the country.)

И, к сожалению, вычеркнуть подобное из памяти невозможно, это не **СМС** на телефоне, которое можно стереть и забыть, это остается с тобой.

(And, unfortunately, it is impossible to cross it out from the memory, it is not an SMS on your phone, which can be erased and forgotten; it stays with you forever).¹⁶

There may be several reasons for this. Firstly, such LAs may be considered cultural borrowings which denote phenomena borrowed at the same time (Haugen, 1950). The Russian translation of the phrase that stands for an LA does not always render the exact meaning, especially in the early stages of a loan's functioning, or the meaning of the Russian phrase may be misleading.¹⁷ For example, if we translate the abbreviation *PR* (public relations) into Russian, without knowing what the term stands for, the phrase *публичные отношения* (public relations) may well be understood as 'somebody's relations in front of the public', which is not exactly what the term *PR* implies.¹⁸ Moreover, this can serve as an example of the clash of cultural values, because demonstrating peculiarities of your life in public is considered inappropriate. The use of the neutral term *пиар* (PR), on the other hand, allows for avoiding such traps.

Another explanation is offered by Rochtchina (2012: 79-80), who believes such cases to be examples of 'double orthography', which uses the principles like those in the Japanese script whereby two alphabets (katakana – for more recent loans, and hiragana – for native Japanese and loans from Chinese) act simultaneously.

Another reason may be pragmatic. LAs are a compact way of representing complex concepts. Word-like versions of English abbreviations are perceived as conceptually whole units, and since the loans of this kind are morphologically non-transparent, the

meaning of the whole is not motivated by the semantics of the components. Therefore, translating the phrases to produce Russian abbreviations will break the cognitive unity of the concept. It will also require more complex morphological-semantic processing and speakers will need knowledge of what the given letters stand for. Considering the tendency for economy in language, the use of a well-established item that is listed in the vocabulary receives priority.

4.6. Limitations on assimilation of loan abbreviations

Derivational potential of LAs can be considered an important factor for their assimilation in Russian. The concept of “derivational potential” is offered by Dyakov (2001), who uses a set of criteria allowing to predict whether a word can or cannot be successfully assimilated in the recipient language. The outline below demonstrates how different factors affect assimilation of LAs in Russian.

4.6.1. The role of phonology

As with loanwords, the most important factor that has an impact on the assimilation of LAs in Russian is conformity with the phonetic/phonological system of the language.¹⁹ For example, the final vowel *-u* in LAs like *сиду* [sidi] ‘CD’ is not typical of Russian singular nouns, which points to the low derivational potential of words with this ending. As a result, abbreviations ending in *-u* are more likely to be used as transplants in their original English spelling as parts of LA+N composites; e.g., *CD-диск* [CD-disk] ‘CD-disk’, *PC-приложение* [PC-prilozheniye] ‘PC application’.

LAs of such kind may still be used as lexemes and derivational bases; however, this use is limited to colloquial speech or professional jargon, e.g. *сидюк* [sidyuk] ‘CD-player/CD-ROM’, *дивидюк* [dividyuk] ‘DVD-player’, *дивидишка* [dividishka] ‘DVD-disk’. Such russified versions are stylistically coloured and belong mainly to expressive vocabulary (Rochtchina, 2012: 78). This effect is enhanced by the use of Russian suffixes that are typical of colloquial speech. For example, references to the derivatives formed on the basis of the LA *сиду* [sidi] ‘CD’, such as *сидюк* [sidyuk] ‘CD-player’, *сидюшник* [sidyushnik] ‘CD-player’ or ‘CD system’, *сидюшный* [sidyushnyi] ‘relating to CD’, in a Google search demonstrate that the words usually occur in IT blogs and forums. The definitions of such words may or may not be given in dictionaries of computer and young people’s jargon:

Мы отцепили **сидюк**, стали цеплять хард, два часа мучились, а комп харда не видит.
(We unhooked the CD-player and began to hook the hard drive; after two painful hours the computer still didn't «see» the hard drive.)

Давай диски в **сидюшнике** поменяем, надо нового чего-нибудь послушать!
(Let's change the discs in the CD system, we need to listen to something new!)

Валялся какой-то кассетно-**сидюшный** в гараже с небольшими колонками типа филипс.
Если интересуется, посмотрю в каком состоянии.

(There was one cassette-CD player in the garage with small speakers, probably Phillips. If you are interested, I'll check its condition.)

These examples of LAs have a limited derivational potential in terms of their phonology. Interestingly, the same holds true for cases that are not limited by the phonological system of Russian. LA-based derivatives like *ФБРовцы* 'FBI officers', *ЦРУшники* 'CIA officers' also demonstrate the presence of negative connotations, which can be tracked through the use of such derivatives in negative contexts in the Russian media. Derivatives like these are mainly used in interviews and blogs rather than bulletins and analytical reports; their use in broadsheets seems to be restricted.

Only very few LAs may be considered successfully assimilated, where the derivatives formed on the basis of the assimilated abbreviation are perceived as the norm of the language and are frequently used in neutral contexts. The loan *пиар* [piar] 'PR' may serve as a good example of this. Vaganova (2005) lists seven derivatives that use the loan as a derivational base; e.g., *пиарить* [piarit'] (ImpV), *пропиарить* [propiarit'] (PerfV), *отпиарить* [otpiarit'] (PerfV), *пиарщик* [piarshchik] (AgtN), *пиармен* [piarmen] (AgtN), *пиарология* [piarologiya] 'PR-ology' (N), *пиаролог* [piarolog] 'PR scientist' (N). The choice of the derivational material is important too. The derivatives with Russian suffixes, i.e. *пиарить* [piarit'] 'to PR', *пропиарить* [propiarit'] 'to have PRed', *отпиарить* [ot-piar-it'] 'to have PRed', *пиарщик* [piarshchik] 'PR specialist', all have somewhat negative connotations. At the same time, derivatives that employ foreign suffixes are stylistically neutral, e.g. *пиармен* [piarmen] 'PR specialist', *пиарология* [piarologiya] 'PRology', *пиаролог* [piarolog] 'PR scientist'.

So, we can see that the trend for the changes in meaning in the process of derivational assimilation of loans (See in Section 3.5 above) can be noticed in the assimilation of LAs too. The fact that this trend is not limited to loanwords is interesting and confirms the suggestion that the use of indigenous Russian derivational material may be responsible for pejoration of the meaning of the concept expressed by the foreign stem, which creates a humorous effect.

4.6.2. The role of prestige/novelty

Another factor that limits the assimilation of the loan is the loss of novelty and/or prestige, and the use of Russian derivational means can be considered a contributor. As pointed out above, derivatives formed on the basis of assimilated abbreviations have a limited range and are more typical of colloquial speech. Thus, it is possible to speculate that with limiting the register in which a loan is used, we may also expect changes in the meaning of the concept that the loan-based derivative denotes. On the other hand, preserving the form of the loan (or the form of the LA), which is possible through the employment of the derivational means of foreign origin (foreign affixes) or the use of the foreign morphological pattern like N+N compounding, allows for keeping its semantic structure and value in the eyes of the speakers intact. The analysis of LA+N sequences that refer to a person working in the IT sphere seems to demonstrate a higher level of respect to the profession compared to the assimilated LA *айтишник* (aitishnik 'a person that has

something to do with IT'), cf. *IT-эксперт* (IT-expert), *IT-профессионал* (IT-professional), *IT-менеджер* (IT-menedger), *IT-любитель* (IT-enthusiast). The russified version *айтишник* sounds colloquial, and for a native speaker conveys the meaning of someone who has some knowledge about/experience with IT, or even an amateur who can help sort out mundane computer-related problems.

So far, we have seen that through derivational assimilation, a loan or an LA (and the concept it denotes) may acquire additional connotations that are not present in the etymon and even lose its prestige and novelty in the eyes of the speakers of the recipient language. Besides the change in the stylistic status of the LA in the process of becoming derivationally assimilated, its general meaning may also change. For example, the loan *ВИП* [vip] 'VIP', which may orthographically be represented as a transplant *VIP* or as a word – *вин*, demonstrates this trend. If we compare the meanings conveyed by the modifying constituents in *виновские места* [vipovskie mesta] (vip.AdjPl seat.Pl) versus *вин-места* [vip-mesta] 'VIP-seats', the meaning conveyed by the derivative in *виновские* can be understood as "cool", "prestigious", which is part of the connotations of the meaning of the original concept expressed by the LA *вин/ВИП/VIP*, but not its core semantic content. At the same time, the meaning of the analytical composite *вин-места* 'VIP seats' is correlated with the semantics of the English etymon, i.e. 'seats for VIPs'.

4.7. Alternative route for assimilation

As can be seen from the discussion above, LAs have the potential to become assimilated in the recipient language, with derivational assimilation being the final stage of a loan becoming russified. After such russification, loans tend to undergo changes in their meaning and are more likely to be used in colloquial speech. This means that in the process of assimilation concepts denoted by loans in general (and LAs in particular) may lose some of the characteristics they were initially borrowed for. The need to preserve the conceptual value of loans seems to be connected with preserving their form. Therefore, the use of LAs in analytical composites, where there is no need for grammatical and derivational assimilation, can be viewed as an alternative route that allows for compensating the aforementioned limitations associated with derivational assimilation. The number of N+N composite rows with LAs in the modifying position supports this argument.

CD

CD-обзор [CD-obzor] 'review of new CDs', *CD-раздел* [CD-razdel] 'CD profile', *CD-чейнджер* [CD-cheindzher] 'CD changer', *CD-проигрыватель* [CD-proigryvatel] 'CD player', *CD-диск* [CD-disk] 'CD disc', *CD-сборник* [CD-sbornik] 'collection of songs on a CD', *CD-архив* [CD-arkhiv] 'CD archive', *CD-плеер* [CD-pleyer] 'CD player', *CD-парад* [CD-parad] 'ranking of CDs', *CD-ресивер* [CD-resiver] 'CD receiver', *CD-бокс* [CD-boks] 'CD box', *CD-носитель* [CD-nositel] 'CD-ROM', *CD-текст* [CD-tekst] 'CD text', *CD-диагноз* [CD-diagnoz] 'diagnostics of the CD drive', *CD-продукция* [CD-produktsiya] 'CDs that are sold'

ВИП/VIP

ВИП/VIP-мнение [VIP-mnenie] ‘opinion of a VIP’, *ВИП/VIP-персона* [VIP-persona] ‘VIP person’, *ВИП/VIP-категория* [VIP-kategoriya] ‘category of products of very high quality’, *ВИП/VIP-клиент* [VIP-klient] ‘VIP client’, *ВИП/VIP-версия* [VIP-versiya] ‘a version of a product that is suitable for VIPs’, *ВИП/VIP-вариант* [VIP-variant] ‘a version of a product that is suitable for VIPs’, *ВИП/VIP-отель* [VIP-otel] ‘hotel where VIPs stay’, *ВИП/VIP-перевозчик* [VIP-perevozhchik] ‘carrier of VIP cargo’, *ВИП/VIP-автомобиль* [VIP-avtomobil] ‘VIP car’, *ВИП/VIP-захоронение* [VIP-zahoronenie] ‘burial place for VIPs’, *ВИП/VIP-участок* [VIP-uchastok] ‘a piece of land for VIPs’, *ВИП/VIP-подписка* [VIP-podpiska] ‘VIP subscription’, *ВИП/VIP-лица* [VIP-litsa] ‘VIPs’, *ВИП/VIP-проход* [VIP-prohod] ‘an aisle or door reserved for the use of VIPs’]

IT

IT-отдел [IT-otdel] ‘IT department’, *IT-департамент* [IT-departament] ‘IT department’, *IT-эксперт* [IT-expert] ‘IT expert’, *IT-технология* [IT-tekhnologiya] ‘IT technology’, *IT-город* [IT-gorod] ‘city for IT specialists and their families’, *IT-бизнес* [IT-biznes] ‘IT business’, *IT-компания* [IT-kompaniya] ‘IT company’, *IT-ипотека* [IT-ipoteka] ‘mortgage with special conditions for IT specialists’, *IT-сфера* [IT-sfera] ‘IT sphere’, *IT-решение* [IT-resheniye] ‘IT solution’, *IT-разработка* [IT-razrabotka] ‘product developed by IT specialists’, *IT-проект* [IT-proekt] ‘IT project’

PR

PR-поддержка [PR-podderzhka] ‘PR support’, *PR-команда* [PR-komanda] ‘PR team’, *PR-акция* [PR-aktsiya] ‘PR campaign’, *PR-тенденция* [PR-tendentsiya] ‘PR tendency’, *PR-вопрос* [PR-vopros] ‘PR issue’, *PR-отдел* [PR-otdel] ‘PR department’, *PR-менеджер* [PR-menedzher] ‘PR manager’, *PR-департамент* [PR-departament] ‘PR department’, *PR-директор* [PR-direktor] ‘PR director’, *PR-специалист* [PR-spetsialist] ‘PR specialist’

SMS

SMS-оповещения [SMS-opoveshcheniya] ‘SMS notification’, *SMS-технология* [SMS-tekhnologiya] ‘SMS technology’, *SMS-рассылка* [SMS-rassylka] ‘SMS mass texting’, *SMS-спам* [SMS-spam] ‘SMS spam’, *SMS-информирование* [SMS-informirovanie] ‘informing via text messaging’, *SMS-уведомление* [SMS-uedomlenie] ‘SMS notification’, *SMS-выписка* [SMS-vypiska] ‘bank statement that can be received via SMS’, *SMS-пароль* [SMS-parol] ‘online access code that is sent via SMS’, *SMS-опрос* [SMS-opros] ‘SMS survey’, *SMS-голосование* [SMS-golosovanie] ‘voting via SMS’, *SMS-сервис* [SMS-servis] ‘SMS service’, *SMS-трафик* [SMS-traffik] ‘SMS traffic’, *SMS-стандарт* [SMS-standart] ‘SMS format’

As can be seen from the examples, some of the N+N composites with LAs could have been borrowed as single units (e.g., *IT-технология* [IT-tekhnologiya] ‘IT technology’) or as loan translations (Haugen, 1950), (e.g., *CD-проигрыватель* [CD-proigryvatel] ‘CD player’). Some have been formed in the recipient language (e.g., *ВИП/VIP-захоронение* [VIP-zahoronenie] ‘burial place for VIPs’).

The discussion in this section aimed to demonstrate that the use of LAs in the modifying position of N+N composites is one way to avoid the limitations outlined above. Since the first component of such sequences is used analytically, the derivational potential of the LAs is not important, as we can see in the case with *CD+N* composites that form an extensive composite row. Preserving the form of the LA also preserves its

original meaning and prestige. In addition, the combination of the LA with another noun avoids the necessity for it to be assimilated on grammatical and derivational levels, which helps to deal with restrictions set by the syntactic system of Russian. The use of LAs in Russian analytical composites can be also considered a special model of hybridisation, in which a whole word is used as a means of accommodating a loan into the recipient language.

4.7. The use of LAs in Russian composites: old vs. new

The last question to touch upon in the frame of this investigation is whether it is right to say that through borrowing of English N+N compounds or LA+N compounds Russian has also borrowed/developed a new word-formation pattern. The use of such structures in the creation of new LA+N composites can be viewed as evidence that the language has accommodated not only the loans themselves but also the word-formation pattern. Rochtchina (2012: 80) claims that the number of LA+N compounds created from Russian material (*PIA-Новости* [RIA-novosti] ‘RIA News’, *БТК банк* [BTK-bank] ‘BTK bank’, *ТВ-программа* [TV-programma] ‘TV programme’, etc.) demonstrates that “... Russian morphosyntax is adopting an English structure.” Even though this research does not entirely support this claim, there is certainly sufficient evidence that Russian is moving towards analytism under the influence of linguistic borrowing. Similar trends have been noticed in other Slavic languages, including Czech (Gazda, 2008), Bulgarian (Bagasheva, 2016), Polish (Sztencel, 2009), as well as Romance languages, e.g. French (Picone, 1996)

Intriguingly, indigenous Russian abbreviations and abbreviations that were borrowed during the Soviet times function somewhat differently from recent LAs. For example, the data in the RNC (n.d.) suggests that Russian-originating abbreviations are rarely used in the formation of composites, but are rather used as post-modifiers in noun phases, e.g. *офицер ЧФ* [ofitser ChF] – ofitser_{Nom.} Chernomorskogo_{Gen.} Flota_{Gen.} ‘Officer of the Black Sea Navy’; *оператор АСУ* [operator ASU] – operator_{Nom.} Avtomatizirovannyh_{Gen.} System_{Gen.} Upravleniya_{Gen.} ‘operator of ASU’).

A possible explanation for this may be that the connection with the motivational base is preserved on a deeper level of mental representation, and the knowledge about how the full phrase is conventionally used (i.e. adnominal genitive construction: N+N_{Gen.}, where LAs take the position of N_{Gen.}) constrains the movement of the analytical element to the pre-modifying position.

The same is true for LAs that have been grammatically and lexically assimilated, are listed and are relatively high-frequency words. For example, LAs like *НАТО* ‘NATO’, *ООН* ‘UNO’, *США* ‘USA’ are not involved in derivational processes and are used in syntactic constructions in a similar way to indigenous Russian loans; e.g., *посол США* ‘USA ambassador’, *заседание ООН* ‘UNO meeting’, *сотрудник НАТО* ‘NATO official’.

Such positional constraints do not seem to be the case for most recently-borrowed abbreviations. One explanation that can be offered here is analogy. The meaning of a foreign abbreviation is often not motivated by the words the abbreviation stands for in

the recipient language, and they are perceived by native speakers as foreign words/parts of words. Since one of the most prominent ways of introducing foreign words into Russian in the last few decades is through nominal compounding, LAs have become used in the pattern in which foreign words often occur—i.e. N+N. A similar phenomenon is registered in other languages. For example, Picone (1996) discusses the possibility of the influence of borrowed English N+N compounds on the overall acceptance of right-headed elliptical N+N sequences in modern French.

It is interesting, however, why older abbreviations are introduced into a syntactic structure by means of a genitival phrase. The study of available sources does not provide a definitive answer, which leaves the way open for speculation. One of the possible explanations may be that abbreviations borrowed earlier, especially during the Soviet rule, were introduced through official media channels. The structure of sequences was carefully checked by the censorship, one of the functions of which was editing the text in accordance with the prescribed rules of Russian. This may have influenced the way such structures are used. After the Soviet state fell apart, and especially with the advent of the Internet, the channels through which the information came into the country changed. The words that started to be borrowed denote the concepts that are relevant for people's everyday lives, which means that they have more communicative value for an individual. This may be considered one of the factors that can explain such change in the use of recent LAs.

It is also possible to speculate that in future, recently introduced LAs may start being used in a way that is more typical of Russian abbreviations: when abbreviation is used in adnominal genitive construction, e.g. *сотрудник ООХ* (UNO officer). It is also possible that their use in compounds may remain preserved in the way it works now; i.e., LA+N. If the pattern passes the test of time, we will have more evidence for claims that the analytical compound formation pattern has been brought into use under the influence of borrowing a large number of English N+N compounds.

5. Conclusion

This paper analysed a sample of analytical N+N composites in Russian, paying special attention to the use of LAs in the leftmost (modifying) position. Based on the analysis, it is possible to suggest that the analytical N+N pattern has become more productive in Russian under the influence of numerous borrowed N+N compounds. However, the straightforward claim that through the borrowing of English N+N compounds Russian also borrowed the analytical N+N word-formation pattern seems premature. The position that takes into consideration the reasons behind the growing productivity of the pattern provides a firmer base for a grounded explanation of the issue. It is important to consider that a superficially similar analytical word-formation pattern already exists in Russian (e.g. *царь-пушка* [tsar'-pushka] 'king cannon'). Since this pattern is unproductive, it is often disregarded in modern research on Russian analytical compounds. It is possible, though, that English N+N compounds have triggered the revival of the pattern, which had previously existed. This revival brings with itself certain limitations; that is, the use of a

recent loanword in the modifying position in the majority of cases and the predominant use of the word of a foreign origin in the head position.

The lack of clarity in terms of the reasons for growing productivity for Russian analytical N+N composites makes it difficult to assess the scale of English influence on the Russian language. Accepting the position that the English N+N pattern was borrowed under the influence of the large number of compounds, the contact between the two languages should be viewed as “more intense” (when some basic and non-basic vocabulary, as well as some structures are borrowed), according to Thomason and Kaufman’s (1988) scale of borrowing. On the other hand, the position that views the use of the N+N pattern as an example of revitalisation of an unproductive one, suggests that the contact between English and Russian is “casual” (with only lexical borrowings). Even though the presented study did not aim to investigate the scale of language contact, further research in this area will be helpful in understanding the reported analytical tendencies in Russian, which are attributed to the influence of English.

The discussion above touched upon the stages of assimilation of a loan in the new linguistic environment with special attention to the effects of derivational assimilation, which involves the use of different derivational means. In the process of derivational assimilation, loans in general and LAs in particular may acquire semantic characteristics which are not registered in their English etymons. Such changes may be attributed to the use of Russian derivational material (hybridisation), which brings in the meaning strongly associated with Russian words and creates a humorous effect.

Russification of the loan is also responsible for downgrading the value of the concept in the eyes of the native speaker of Russian, and sometimes pejoration of meaning even in cases with the use of stylistically neutral affixes. Further research on the influence of different derivational means in the process of the assimilation of the loan will allow for a better understanding of native speaker attitudes to the borrowed concepts as well as on the process of linguistic borrowing into Russian.

The need to preserve the conceptual value of the loans and/or LAs in the eyes of the native speakers of the recipient language seems to be strongly associated with the preservation of their structural form. The presented analysis suggests that the use of loan abbreviations in analytical composites, where there is no need for a loan to be assimilated (even on the orthographic level), can be viewed as a logical way to keep the formal characteristics of the loan, its semantic content and conceptual value intact. The use of LAs in the modifying position of N+N composites also avoids the morphosyntactic restrictions of the Russian language system and compensates for the low derivational potential of some LAs. At this stage, we can see that the analytical N+N pattern serves as a catalyst that brings new concepts into the system of the Russian language and is becoming more productive.

The use of Krysin’s (1975) framework is determined by the necessity for looking at assimilation of loans as a gradual process, and views loans from the viewpoint of their degree of assimilation. A study that would explore the possibilities of categorising borrowings into Russian using more widely accepted classifications, e.g. Thomason and Kaufman’s (1988) borrowing scale, Haugen’s (1950) classification, Matras and Sakel’s

(2007) framework of MAT (matter) and PAT (pattern) integration of loans, Carstensen and Busse's (1994) classification of forms of linguistic transfer, etc. in order to contribute to the development of a typology of linguistic borrowing, would be valuable.

Another limitation of the current study is that it is clustered around a limited number of semantic domains of language use with the main focus on the use of the analytical N+N pattern in technical language and business communication. Expanding the analysis beyond these areas and looking at the attitudes of native Russian speakers of different age groups towards this pattern will allow for employing a usage-based approach and will provide further insights into native speaker cognition.

Notes

1. Some of Haugen's (1950) categories are used where appropriate in this discussion.
2. This approach to treatment of the constituents of a Russian complex word raises questions connected with the terminology. For example, it is not clear whether understanding of the first constituent as a morpheme is determined by its representation by the stem or whether something else is going on. However, for reasons of space, and because the matters of terminology are not crucial for this investigation, I will refrain from discussing these here.
3. This assumption is rather simplistic for the purposes of the discussion. The criteria commonly used for differentiating between a compound word and a phrase in Russian also include structural (combination of two or more lexical centres, i.e. roots or stems, in a prescribed order), grammatical (the loss of grammatical markers by the first element), and semantic (unified meaning).
4. It has been pointed out by one of the reviewers that *zhar-ptitsa* may be a calque. This possibility cannot be excluded, since there are formations describing somewhat similar concepts in other Slavic languages, e.g. Czech *pták ohnivák* 'bird-fire' and Slovenian *zlata ptic* 'golden bird'. The analysed etymological sources do not provide a definitive answer to the question of the origin of this N+N sequence, referring it to ancient Russian mythology. This example also supports the argument that classifications of loans that involve etymological principles may be somewhat misleading since it can be hard to draw a borderline between an indigenous word and a loan in the absence of phonological and/or morphological similarity.
5. Examples are taken from the RNC (n.d.).
6. Interestingly, in the RNC transplants (items using the script of the donor language) are tagged as barbarisms, whereas once the Cyrillic script is used, they acquire the status of a loanword.
7. Rochtchina (2012: 79) notes the overall tendency to preserve the English orthography in abbreviations that are related to information technology, e.g. DOS, IBM. She also points out that Cyrillic forms of abbreviations—e.g. *Бу-бу-су* (BBC), *МІТЗ* (MP3)—often revert to their Latin alphabetic forms.
8. Cp. Eng.: *parking* is motivated by the derivational base *park* and has the meaning 'bring (a vehicle that one is driving) to a halt and leave it temporarily, typically in a car park or by the side of the road' ("Parking", n.d., n.p.).
9. Matiello (2013) provides a detailed analysis of the issues that are associated with abbreviations in modern linguistics, including problems of definition, classification, irregularities in formation, etc.
10. Examples are taken from the RNC (n.d.).

11. This phenomenon is also referred to as RAS syndrome or PNS syndrome and is not limited to loan abbreviations. The phenomenon is widespread and supports the suggestion that abbreviations are perceived as words in general, e.g. English *ATM machine* [[Automated Teller Machine] machine], *PDF format* [[Portable Document Format] format].

12. The RNC (n.d.).

13. It is hard to say whether there is a tendency for the transplants to become graphically assimilated. The texts in the RNC contain at least ten times the number of examples of use of the LA *VIP* in the English orthography as compared to its number of uses in the Russian spelling *вун/ВУИИ*. The dates of the texts do not seem to matter either. Considering the limited size of the RNC, this cannot be considered representative of the overall tendencies, because a search in Google demonstrates that the use of the LA in the Russian orthography is quite common (over 200,000 hits).

14. For example, in the cases of *ЦПВ* (Eng.: CIA) and *ФБР* (Eng.: FBI), the syntactic structure of the source phrases in English and Russian do not match, which is not the case for *ПК* (Eng.: PC). For a more detailed description of abbreviation calques and differences in their structure in Russian see Baranova (2010).

15. Examples are taken from RNC (n.d.).

16. Examples are taken from RNC (n.d.).

17. It is important to point out that the tendency for russification of English abbreviations is becoming stronger. However, some abbreviations can be considered successful translations, e.g. *США* (USA), *ФБР* (FBI).

18. This could have been the case for English too, before the expression *public relations* started to be used as a term.

19. Other criteria include the length of the loan, presence of semantic competition, conformity with the grammatical system of the language and importance of the concept to the speakers. Dyakov's (2001) understanding of derivational potential seems to be strongly correlated with Alain Rey's comments on neology in French (see Rey, 1995: 80-82).

References

- Bagasheva, Alexandra (2016): "On [N₁N₂] constructions and word-formation in Bulgarian". In L. Körtvélyessy et al., eds., *Word-formation Across Languages*. Newcastle upon Tyne: Cambridge Scholars Publishing, 1-28.
- Baranova, Lyudmila (2010): "Kalkirovanie kak sposob zaimstvovaniya inoyazychnykh abbreviatur v russkom yazyke". *Сучасні дослідження з іноземної філології*, http://nbuv.gov.ua/UJRN/Sdzif_2010_8_6 [Accessed 25 October 2017]
- Bauer, Laurie (2003): *Introducing linguistic morphology*. Washington, DC: Georgetown U Press.
- Broselow, E. (1992): "Loan phonology". In W. Bright, ed., *International Encyclopedia of Linguistics: EXPL – MOOT*. Oxford: Oxford University Press, 200-201.
- Busse, Ulrich, and Manfred Görlach (2002): "German". In M. Görlach, ed., *English in Europe*. Oxford: Oxford University Press, 13-36.
- Carstensen, Broder, and Ulrich Busse (1994): *Anglizismen-Wörterbuch. Der Einfl uss des Englischen auf den deutschen Wortschatz nach 1945*. Berlin, New York: Mouton de Gruyter.
- Chernykh, P.Y. (1999): *Istoriko-etimologicheskyyi slovar' covremennogo russkogo yazyka*. 3rd ed., Moscow: Russkiy yazyk.

- Corbett, Greville (2004): "The Russian adjective: a pervasive yet elusive category". In R.M.W. Dixon and A. Y. Aikhenvald, eds., *Adjective Classes: A Cross-Linguistic Typology*. Oxford: Oxford University Press, 199-222.
- Crystal, David (2008): *A Dictionary of Linguistics and Phonetics*. 6th ed., Oxford: Blackwell Publishing.
- Dimitrova, Ludmila, et al. (2011): "Bulgarian-Polish-Lithuanian Corpus – Recent progress and application". In D. Majchráková and R. Garabí, eds., *Natural Language Processing, Multilinguality. Sixth International Conference Modra, Slovakia, 20–21 October 2011 Proceedings*. Slovak National Corpus L. Štúr Institute of Linguistics Slovak Academy of Sciences Bratislava, Slovakia, 44-50.
- Dyakov, Anatoly (2001): *Derivatsionnaya integratsiya anglitsizmov v russkom yazyke kontsa XX veka v funktsional'nom aspekte*. PhD dissertation, Novosibirskiy Gosudarstvennyi Pedagogicheskyy Universitet.
- Fasmer, Max (1986): *Etimologicheskyy slovar' russkogo yazyka*. 2nd ed., Moscow: Progress.
- Fischer, Roswitha (2008): "Introduction: Studying anglicisms". In R. Fischer and H. Pułaczewska, eds., *Anglicisms in Europe: Linguistic Diversity in a Global Context*. Newcastle upon Tyne: Cambridge Scholars Publishing, 1-14.
- Gazda, Jiří (2008): "Rost analitizma (i agglutinativnosti) kak odna iz konvergentnykh tendencij v razvitii sovremennykh slavyanskikh jazykov (na materiale novykh slozhnykh naimenovanij v russkom i cheshskom jazykach)". *Izuchenije slavyanskikh jazykov, literatur i kultur kak inoslavjanskikh i inostrannykh. Sbornik dokladov*. Belgrad: Slavisticheskoe obshchestvo Serbii, 118-123
- Gómez Capuz, Juan (1997): "Towards a typological classification of linguistic borrowing (illustrated by Anglicisms in Romance Languages)". *Revista Alicantina de Estudios Ingleses*, 10: 81-94. doi: 10.14198/raei.1997.10.08. [Accessed 25 October 2017]
- Haspelmath, Martin, and Uri Tadmor (2009): "The loanword typology project and the world loanword database". In M. Haspelmath and U. Tadmor, eds., *Loanwords in the World's Languages: A Comparative Handbook*. Berlin: de Gruyter, 1-34.
- Haspelmath, Martin (2009): "Lexical borrowing: Concepts and issues". In M. Haspelmath and U. Tadmor, eds., *Loanwords in the World's Languages: A Comparative Handbook*. Berlin: de Gruyter, 35-54.
- Haugen, Einar (1950): "The analysis of linguistic borrowing". *Language*, 26(2): 210-231.
- Hyman, Larry (1970): "The role of borrowings in the justification of phonological grammars". *Studies in African Linguistics*, 1(1): 1-48.
- Karaulov, Yury, ed. (1998): *Russkiy yazyk: entsiklopedia*. 2nd ed., Moscow: Drofa.
- Kimyagarova, Roza (1968): "Tipy i vidy adaptatsii zaimstvovannoi leksiki v russkom yazyke novogo vremeni (XVII-XX vv.)". *Vestnik Moskovskogo Universiteta. Seriya 9. Filologiya*, 6: 69-78.
- Komlev, Nikolai (2006): *Slovar' inostrannykh slov*. Moscow: EKSMO-Press.
- Kotelova, Nadezhda, ed. (1984): *Novye slova i znacheniya. Slovar'-spravochnik po materialam pressy i literatury 70-h godov*. Moscow: Russkiy yazyk.
- Krysin, Leonid (1975): "Stupeni morfemnoi chlenimosti inoyazychnykh slov". In E. Zemskaya, ed., *Razvitie sovremennogo russkogo yazyka. Slovoobrazovanie. Chlenimost' slova*. Moscow: Prosveshchenie, 227-231.
- Matiello, Elisa. (2013): *Extra-Grammatical Morphology in English. Abbreviations, Blends, Reduplicatives and Related Phenomena*. Berlin: Mouton De Gruyter.

- Matras, Yaron, and Jeanette Sakel (2007): "Investigating the mechanisms of pattern replication in language convergence". *Studies in Language*, 31(4): 829-865.
- Nagano, Akiko, and Masaharu Shimada (2016): "Language contact between English and Japanese and the borrowing of left-headed nominal modification construction". *ESSE 2016: Book of Abstracts*: 51. www.esse2016.org/downloads/abstracts.pdf
- Panov, M. V. (1999): *Pozicijonnaja morfologija ruskogo jazyka*. Moscow: Nauka, 1999.
- "Parking". n.d. In *Oxford Dictionaries*, en.oxforddictionaries.com/definition/parking.
- Picone, Michael D. (1996): *Anglicisms, neologisms and dynamic French*. Amsterdam / Philadelphia: John Benjamins Publishing Company.
- Rey, Alain (1995): *Essays on Terminology*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Rochtchina, Julia (2012): "Morphology and lexicology interface. Latest Russian neologisms: The next step towards analytism?" In V. Makarova, ed., *Russian Language Studies in North America: New Perspectives from Theoretical and Applied Linguistics*. London, New York: Anthem Press, 71-84.
- Russian National Corpus. N.p., n.d. Available at ruscorpora.ru/
- Ryazanova-Clarke, Larissa, and Terence Wade (1999): *The Russian Language Today*. New York: Routledge.
- Shvedova, Yelena, ed. (1980): *Russkaya grammatika*, vol.1. Moscow: Nauka.
- Štekauer, Pavol, et al. (2008): *Universals, Tendencies and Typology in Word-Formation: A Cross-Linguistic Study*. Josef Šafárik University.
- Sztencel, Magdalena (2009): "Boundaries crossed: The influence of English on modern Polish". *E-Pisteme*, 2(1): 3-17.
- Thomason, Sarah Grey, and Terrence Kaufman (1988): *Language Contact, Creolization and Genetic Linguistics*. Berkeley: University of California Press.
- Tikhonov, Alexandr, ed. (2014): *Entsiklopedičeskij slovar'-spravočnik lingvističeskij terminov i ponyatij: Russkij jazyk*, T.1. 2nd ed., Moscow: Flinta.
- Townsend, Charles (1968): *Russian Word-formation*. New York: McGraw Hill.
- Vaganova, Natalya. (2005): *Sovremennye zaimstvovaniya iz anglyiskogo jazyka: Semantiko-slovoobrazovatelnyj aspekt: Na materiale anglicizmov kontsa XX – nachala XXI v. v sovremennom ruskom jazyke*. PhD dissertation, Nizhegorodskij Gosudarstvennyj Universitet.
- Vendina, Tatyana (1999): "Yazykovoje soznanie i metody ego issledovaniya." *Vestnik MGU. Ser. 19. Lingvistika i mezhkul'turnaya kommunikacija*, 4: 15-31.
- Volodarskaya, Emma (2002): "Zaimstvovanie kak otrazhenie rusko-anglyiskih kontaktov". *Voprosy jazykoznanija*, 4: 96-118.

MASTERS PROGRAMME IN ENGLISH AND SPANISH FOR SPECIFIC PURPOSES Curso 2017-2018

PREINSCRIPCIÓN, MATRÍCULA Y TASAS

Preinscripción: solicitud mediante formulario electrónico desde mayo hasta septiembre, en varios periodos, gestionada por el Centro de Formación Continua. Una vez admitido/a se te indicarán los pasos a seguir, para que puedas matricularte on-line en tu Facultad o Escuela a través de UACloud CV.

Tasas: las tasas de matrícula las establece anualmente la Generalitat Valenciana, mediante decreto del Consell.

* info: <http://web.ua.es/es/continua/masteres-universitarios.html>

INGLÉS Y ESPAÑOL PARA FINES ESPECÍFICOS

INFORMACIÓN Y CONTACTO

- Facultad de Filosofía y Letras
- facu.iletres@ua.es
- iulma@ua.es
- 955 903 448
- 955 903 449
- web.ua.es/es/masteres-oficiales.html

Universitat d'Alacant
Universidad de Alicante

IULMA
Instituto Universitario de Lingüística y Literatura
de la Universidad de Alicante
Instituto Universitario de Lingüística y Literatura
de la Universidad de Alicante

INFORMATION

University of Alicante – Faculty of Arts
<http://web.ua.es/es/masteres-oficiales.html>

Registration info:

Lifelong Training Centre (ContinUA)

<http://web.ua.es/es/continua/masteres-universitarios.html/>



Universitat d'Alacant
Universidad de Alicante

Present-day Spanish fashion lexicon dresses up in English¹

Beatriz Rodríguez Arrizabalaga
Universidad de Huelva
arrizaba@dfing.uhu.es

ABSTRACT

Lexical borrowings can be regarded as one of the clearest and most direct consequences of any language contact situation. However, not all the borrowings that enter a language are alike. Since their entrance in a given language is motivated by different reasons, two general kinds of borrowings must be distinguished: necessary borrowings which name ideas and concepts for which the recipient language does not have any equivalent term; and superfluous borrowings which, on the contrary, refer to realities for which the recipient language already has equivalent terms. This paper focuses on the latter type. Specifically, it presents a diachronic corpus-based analysis of 14 English fashion terms with a clear Spanish lexical counterpart — *blazer*/'chaqueta', *celebrity*/'famoso', *clutch*/'bolso de mano', *cool*/'de moda', *fashion*/'moda', *fashionable*/'de moda', *fashionista*/'adicto a la moda', *jeans*/'vaqueros', *nude*/'color carne', *photocall*/'sesión de fotos', *shorts*/'pantalones cortos', *sporty*/'deportivo', *trench*/'trinchera, gabardina', and *trendy*/'moderno' — in four Spanish corpora: the *Corpus del Español*, and the *CORDE*, *CREA* and *CORPES XXI* corpora. My objectives are twofold: firstly, to demonstrate to what extent these unnecessary Anglicisms are increasingly becoming part of the everyday contemporary Peninsular Spanish fashion lexicon; and secondly, to account for the three reasons that underlie their alleged constant entrance in twenty-first century Peninsular Spanish: (i) globalization and the impact of English on Spanish; (ii) the highly visible presence of English in the field of advertising; (iii) and the selling power of English.

Key words: anglicism, borrowing, advertising, Peninsular Spanish, diachronic corpus-based analysis



1. Introduction

One of the clearest and most direct consequences of any language contact situation is borrowings. These are a “natural” and “omnipresent” phenomenon, as García Morales et al. (2016: 33) and Gómez Cápuz (2004: 9), respectively, describe it, that is to be understood as the exchange of linguistic elements, usually lexical items, between two linguistic communities or two historical languages.

Therefore, as stated in Cronin (2003: 171), “the more we enquire into the particular circumstances of particular languages, the more we discover the multiple traces of borrowings, foreign contacts and wholesale appropriation”. Given, for instance, the historical, cultural, and socio-linguistic evolution of Spanish, as briefly suggested in the following quotation, we can understand why the Spanish language has always been so prone to adopting borrowings, if for different reasons:²

En este sentido, la lengua española, pese a su creciente importancia demográfica y cultural, ha sido tradicionalmente importadora de préstamos: en otras épocas como reflejo de una encrucijada de pueblos y culturas; en los dos últimos siglos, como reflejo de una dependencia técnica y científica de otras culturas occidentales más avanzadas, en especial la angloamericana. (Gómez Capuz, 2004: 9)

Due to the “worldliness of English”, as Pennycook (1994: 33) calls it, or in like terms, the privileged linguistic status that English has over the rest of the world languages as a result of the globalization process we are experiencing in the twenty-first century (cf. Edwards, 1994; Brennan, 1997; Pennycook, 2001; Hjarvard, 2004; Pulcini et al., 2012), English has changed the status of many of the world languages, among which Hjarvard (2004: 76), for example, places different European and Afro-Asian tongues of a very diverse linguistic origin —Germanic (German), Romance (French and Spanish), Slavic (Russian) and Semitic (Arabic):

Over the past two or three decades, English has come to occupy a singular position among languages. Previously only one among several dominant European languages, on a par with French or Spanish, it is today a *world language*, the language people use whenever they wish to communicate with others outside their own linguistic community. English has become the *lingua franca* of the global network [...] As English has moved toward paramountcy, the status of the other principal languages has changed. Even though they are spoken by more people today than ever before, they have been demoted, degraded in relation to English. Today, French, Spanish, Arabic, German, Russian, etc., more or less have the status of regional languages, national languages that can be used beyond their national frontiers. But, they are losing their currency as the language of international communication, formal and informal: both in political and commercial contexts and in intercultural exchanges, as bridges between people who cross cultural frontiers or who like to enrich their lives with media products from abroad.

Therefore, as a result of being the unanimously recognized contemporary *lingua franca* in international communication (cf. De Mooij, 1994: 5; Edwards, 1994:

41; Hjarvard, 2004: 76; Montes Fernández, 2006: 217; Schmidt and Diemer, 2015: 11), the frequent attestation of English in Spanish should not be considered an exceptional fact.

What may strike the reader as surprising, nevertheless, is that, although it is the most recent one, the Anglo-American influence in Spanish has been, as Gómez Capuz (2004: 24-25) describes it, the most “booming” and “overwhelming” foreign trace in the Spanish language since the eighteenth century onwards, pervading not only its lexicon, but also all the remaining linguistic levels: morphology, syntax, semantics and phraseology. In particular, its presence in the lexicon constitutes for Lorenzo (1996: 18), among others, a problem that is becoming worse with the passing of time due mainly to the extremely disproportionate number of English lexical borrowings —many of them unnecessary— attested in contemporary Spanish.³

[...] quedan por despejar bastantes problemas. Creo que el principal es cuantitativo y consiste en la abrumadora penetración, numéricamente sin precedentes, de términos ingleses discutiblemente necesarios.

It is precisely on superfluous Anglicisms that the present paper focuses. In order to validate the previous hypothesis, already set forth by Lorenzo (1996: 18) near the end of the twentieth century, the present paper seeks to demonstrate, first, to what extent gratuitous Anglicisms are more present in contemporary Spanish than they have ever been, and secondly, to explain the three main reasons that, in my view, account for their alleged incessant and striking entrance in present-day Peninsular Spanish. Specifically, it offers a diachronic corpus-based study, covering from the tenth to the twenty-first centuries —namely, up to 2015. With superfluous Anglicisms in view, the paper analyzes the actual use and frequency of occurrence of 14 English fashion terms in the different periods of this particular variety of Spanish. The terms are unnecessary, given that they all, except for *nude* ‘color carne’, have a clearly equivalent original Spanish term, according to the bilingual dictionaries checked: *blazer* ‘chaqueta (deportiva, de colegio)’, *celebrity* ‘famoso’, *clutch* ‘bolso de mano’, *cool* ‘de moda’, *fashion* ‘moda’, *fashionable* ‘de moda’, *fashionista* ‘adicto a la moda’, *jeans* ‘vaqueros’, *photocall* ‘sesión de fotos (para la prensa)’, *shorts* ‘pantalones cortos’, *sporty* ‘deportivo’, *trench* ‘trinchera, gabardina’ and *trendy* ‘moderno’.⁴ The corpora chosen to carry out my analysis are, on the other hand, the *Corpus del Español*, created by Davies (2002), and three of the Spanish Royal Academy corpora: (i) the *Corpus Diacrónico del Español* (CORDE); (ii) the annotated version of the *Corpus de Referencia del Español Actual* (CREA); (iii) and the *Corpus del Español del Siglo XXI* (CORPES XXI).

The paper is structured as follows. In addition to the introductory section (section 1), where the topic of my study, its objectives and the methodology used are outlined, it contains four sections essential to frame and contextualize the study. The first one (section 2) elucidates three of the main problems inherent to the complex phenomenon of lexical borrowings known as “Spanish Anglicisms”:⁵ namely, (i) the terminological confusion that exists around the term; (ii) the dubious or non-English origin of some Anglicisms; (iii) and finally, the well-known distinction between necessary and

superfluous Anglicisms. The second section (section 3) deals with the interconnection that exists between the current process of globalization, the different contemporary mass media, with special emphasis on advertising, and the English language. The third section in this part (section 4) focuses, in turn, on advertising language, which I firmly believe to be in large measure responsible for the widespread use of English in the rest of the world languages. Specifically, three of its distinguishing features will be accounted for: (i) its careful planning; (ii) its connotative load; (iii) and its lexical creativity and innovation. Finally, in section 5, the role that English plays in international advertising, due mainly to its selling power, will be described. Afterwards, in section 6 a detailed description of the data analyzed, the methodology employed, and the findings obtained in my corpus-based study is offered. Special emphasis will be given here, on the one hand, to the treatment —presence/absence— that each of the Anglicisms at issue has received in the lexicographical sources consulted and, on the other, to their frequency of occurrence in the different periods of Peninsular Spanish examined. The paper closes with a concluding section (section 7) where, first, the most significant issues raised are gathered together and, to finish, some suggestions for further research are also made.

2. On Linguistic Borrowings. The Case of Spanish Anglicisms

Despite having always attracted the attention of linguists and speakers, as Gómez Capuz (2004: 11) remarks, linguistic borrowings have never occupied a coherent and central position in the different linguistic subfields, being a marginal topic both in traditional branches like Dialectology and Historical Linguistics and in more modern ones like, for instance, Sociolinguistics. In any case, as pointed out earlier, of all the kinds of borrowings that enter a language, those concerning its lexicon have been undoubtedly the most widely researched. This is so because the lexical level, as Novotná (2007: 6) states, is the part of the language that changes most quickly, since new words appear constantly and, conversely, others fall into disuse or disappear completely.

Since lexical borrowings are necessarily adopted into a language when it lacks an original equivalent term to refer to the extralinguistic reality which the borrowing denotes (cf. Edwards, 1994: 76; Urrutia Cárdenas, 2001: 13; Schmidt and Diemer, 2015: 15), they have to be interpreted as the linguistic consequence of a sociocultural contact between different communities. Although on account of this reason Weinreich (1953) proposed a mixed and broad approach for their study, linguistic at the same time as sociocultural, until very recently the research on lexical borrowings has only been approached either linguistically or sociologically. It has been framed, in fact, in one of the two large and different “schools” or “traditions”, as Gómez Capuz (2004: 13) calls them, where the analysis of such a complex phenomenon originated: on the one hand, the North-American tradition (cf. Bloomfield, 1933; Haugen, 1950; Weinreich, 1953), which, socially oriented, is mainly interested in examining the influence of American English in the diverse European linguistic communities of immigrants located in USA; and on the other, the European school (cf. Betz, 1949; Deroy, 1956; Klajn, 1972), which, being, in turn, more linguistically oriented, focuses on the phenomenon of lexical borrowings between

European languages of a similar prestige and status, like, for instance, English, French, Italian, German and Spanish.

The existence of these two schools, with such diverse and distinct aims, objectives and orientation, has provided many and varied terms for the phenomenon at issue, creating some terminological confusion. Notice in this regard that, whereas the main representatives of the North-American tradition, Bloomfield (1933) and Weinreich (1953), among others, refer, respectively, to borrowings as “intimate borrowings” and “interference”, the term most commonly used within the adherents of the European School (cf. Betz, 1949; Deroy, 1956; Klajn, 1972) is that of “cultural borrowings”. There are, furthermore, some other labels to name borrowings, such as Crystal’s (1997: 22) “loan words” and Newmark’s (1988: 82) concept of “transference”, among others, which make the situation even more complex.⁶

In particular, in the Spanish linguistic tradition, most of the studies carried out on lexical borrowings concern foreign terms coming from English which, due to their linguistic origin, are called Anglicisms (cf. Alfaro, 1948; Fernández García, 1972; Pratt, 1980; Lorenzo, 1996; Medina López, 1996). Based on the most traditional formalist typologies of this particular sociolinguistic and cultural phenomenon, proposed mainly by Betz (1949), Deroy (1956) and Klajn (1972), among others, they usually distinguish, in general terms, though with different names, the three following categories of Spanish Anglicisms:

- (i) Those terms, known as “Patent Anglicisms” (cf. Pratt, 1980; Lorenzo, 1996), which are easily recognized as English, either because they maintain unchanged in Spanish their original English form, like, for example, *hippy* and *ranking* —“raw Anglicisms” (*Anglicismos crudos*) for Lorenzo (1996) and “foreign words” (*extranjerismos*) for Alfaro (1948)— or because, despite their adaptation to Spanish orthographic norms, as can be appreciated in *boicot* and *güisqui*, some clear English trace can be found in them. These are generally referred to as “assimilated Anglicisms” (*Anglicismos asimilados*) by Lorenzo (1996) and “barbarisms” (*barbarismos*) by Alfaro (1948).
- (ii) Those original Spanish terms in form, like, for instance, *romance*, that adopt a foreign meaning in their use: i.e. “amoríos”. Usually included in the general class of “semantic calques” (*calcos semánticos*), they belong to the class of “non-patent Anglicisms” (*Anglicismos no patentes*) in Pratt’s (1980) classification, and to that of “univerbal calques” (*calcos unimembres*) in Lorenzo’s (1996: 492) system.
- (iii) And finally, those lexical calques of English compound words, usually classified as multiverbal so as to be distinguished from those in the previous group, which as Lorenzo (1996: 559) and Gómez Capuz (2004: 40) observe, are literal word-by-word translations into Spanish of English compound terms, no matter their written form. Some examples of this kind, taken from Gómez Capuz (2004: 40), are *basketball* > *baloncesto*, *garden city* > *ciudad jardín*, and *welfare state* > *estado del bienestar*.

Despite their interesting and significant contribution to the study of the universal phenomenon of borrowing,⁷ these works on Spanish Anglicisms entail two main

problems that deserve special attention here. On the one hand, the dubious or non-English origin that, according to Lorenzo (1996: 22-23; 29), some of the terms included in their classifications, like, for instance, *vermut*, *esquí*, and *déficit*, have, due mainly to the “omnivorous” capacity of English to “digest” and “assimilate” all the lexical material that can be taken advantage of, no matter its provenance.⁸ And on the other, the debate about the well-known dichotomy between necessary and superfluous Spanish Anglicisms (cf. Lorenzo, 1996: 18; Urrutia Cárdenas, 2001: 13; Schmidt and Diemer, 2015: 15; Rodríguez González, 2016: 8, among others).

Though it is true, as Lorenzo (1996: 22-23; 29) acknowledges, that many of the foreign terms classified as Anglicisms do not come directly from English, they cannot be disregarded and overlooked in any study on English borrowings owing to the crucial role that English plays as an intermediate language in their spread over the world languages. A good example here is López Morales (1987: 303) who, drawing on the frequent distinction established in the literature between direct and indirect borrowings to account for their different etymology (cf. Katamba, 2015: 136), includes both kinds in his broad definition of Anglicisms:⁹

[...] no sólo palabras que proceden del inglés, independientemente de que sean ya generales en el español y de que hayan sido aceptadas por la Academia, sino también aquellas que proceden de otras lenguas, pero que han entrado al español a través del inglés.

In addition, as regards the attitudes towards the entrance of Anglicisms into Spanish, three different positions —the so-called “purist”, “moderate”, and “integrative” points of view, following Schmidt and Diemer (2015: 15-16)— are to be identified. According to the purist position, necessary Anglicisms are a conscious linguistic and cultural enrichment because they comprise those English terms that are really needed in the recipient language, where no equivalent original terms exist to refer to the extralinguistic reality which they designate; usually cultural products, new “technological developments” and the results of “the reorganization of the global economic space” (Cronin, 2003: 121). Their use is, therefore, as Pennycook (1994: 9) states, “natural, neutral and beneficial”. Gratuitous Anglicisms, on the contrary, should be avoided or, at least, reconsidered since they can be replaced by original words of the lexical inventory of the recipient language. In Urrutia Cárdenas’s (2001: 13) view, their use is, therefore, motivated either by some kind of linguistic incompetence or cultural snobbery, which, for Rodríguez González (2016: 8), moreover, entails a sign of foreignness and modernity. From the purist point of view, consequently, not every single foreign word can be freely accepted in the language, as Urrutia Cárdenas (2001: 12) asserts, without thorough judgment and thought.¹⁰

For the “moderate” and “integrative” attitudes, however, such a radical distinction between necessary and unnecessary Anglicisms should be discarded. For the former, because the English influence in Spanish, by “filling up gaps in the lexicon” and “bringing in new cultural concepts”, is always “enriching” or at least “non-threatening”; and for the latter, which “neutrally describes language contact, as well as its phenomena and change”,

because Anglicisms are to be “documented, but not criticized or restricted” (Schmidt and Diemer, 2015: 15-16).

3. Globalization, the Media and English

As Geertz (1986: 121) remarks, we are living “in the midst of an enormous collage”, where, as stated in Cronin (2003: 169), “peoples, culinary and musical traditions, forms of dress, furnishings, and iconographies are all juxtaposed and thrown together”. There is no doubt, therefore, that we are living in a global world and era (cf. Silverstone, 1999: 106; Geertz, 1986: 121; Cronin, 2003: 169; Williams, 2011: 28). Nevertheless, understanding what globalization means is not easy because, on the one hand, there is little consensus about what it is (cf. Cronin, 2003; MacGillivray, 2006), and on the other, as put forward in Held et al. (1999: 436), because there is “no single coherent theory of globalization and the empirical data generated to assess the impact of global change is limited and contradictory”.

Be that as it may, with these two ideas in mind, globalization can be briefly and generally described, following Robertson (1992: 8), as “the compression of the world”. Therefore, since “English is in the world and the world is in English”, as Pennycook (2001: 78) observes, it is not surprising that over the last thirty years the English language has become the unanimously recognized global *lingua franca* (cf. De Mooij, 1994: 5; Edwards, 1994: 41; Hjarvard, 2004: 76; Montes Fernández, 2006: 217; Schmidt and Diemer, 2015: 11).

The mass media in the twenty-first century world has indeed greatly contributed to the promotion of English as the contemporary international language for several reasons:¹¹ first, because, since it is “assumed to have created a ‘global culture’”, as De Mooij (1994: 4) remarks, there is no globalization without media and communications (cf. Rantanen, 2002: 1; 2005: 4; Kelly-Holmes, 2005: 143); secondly, because the mass media pervades our lives to such a great extent that human beings have become highly dependent on it (cf. Silverstone, 1999: 1-2; Gripsrud, 2002: 3; Jaworski, 2007: 271); in the third place, because, connecting “us to the world outside” and reminding “us that we are members of a society and a world with [...] many places and levels”, in Gripsrud’s (2002: 4) words, the different mass media have such an enormous and powerful socio-cultural impact on our lives (cf. De Mooij, 1994: 121; Urrutia Cárdenas, 2001: 12; Jaworski, 2007: 271) that for Urrutia Cárdenas (2001: 12), in particular, it is even greater than that of educational action:

Los medios de comunicación social (radio, prensa, televisión e internet) son también un importantísimo vehículo de cultura y de formación lingüística. La influencia social de estos medios es extraordinaria. Sin duda, tienen mayor influencia que la acción escolar.

And finally, because the language of international mass media and communications is, undoubtedly, English.

Within the different contemporary mass media, advertising plays a pivotal role in the spread of English to the other languages of the world, since it is, without any doubt, the means of communication where the aforementioned features most clearly converge.¹² In fact, the role which advertising plays in the increasing globalization of the communications industries cannot be denied (cf. Englis 1994: vii; Lester, 1994: 4; De Mooij, 1994: 5; Montes Fernández, 2006). A simple look around us is, indeed, more than enough to perceive the globalized environment of contemporary worldwide advertising:

Take a look around you: you get up in the morning, put on your Italian shoes, get into your German car, go to your office where you use an American computer, and in the evening you watch your Japanese television set, eating a piece of Dutch cheese and drinking a glass of French wine or Scotch whisky. This is the environment of worldwide advertising: the integration of domestic, international, multinational and global or transnational business. (De Mooij, 1994: 5)

On the other hand, as Cook (2008a: 2) explains, its ubiquity and influence in our lives, not only reflecting but also creating culture (cf. Cook, 1992: 5; 2008b: 237; Lester, 1994: 7; Ferraz Martínez, 1995: 10; Fairclough, 2008: 219), are two other incontestable facts which cannot be questioned either:

The contemporary advertisement is everywhere in our lives, jumping out at us from inside the older genres it has come to dominate. It colonises our screens, interrupts our entertainments, punctuates our news, plasters our walls, lines our roadsides, mingles with public information and decorates almost every object we buy. [...] and it frames, represents—arguably even determines—both the broad social issues and the narrow personal agendas of our lives.

And finally, its adoption of English as *lingua franca*, as will be elucidated in section 5, also leaves no room for doubt (cf. De Mooij, 1994: 5; Bathia, 2008: 166).

4. The Language of Advertising

Smith (1982: 190) has described the language of advertising as a “functional dialect” since, as explained in Kelly-Holmes (2005: 8), it is used for a particular purpose. Thus, advertising language has to be understood as a specific linguistic variety, different from normal and everyday expression, which, following Leech (2008: 176), is characterized in terms of its attention value, readability (or ‘listenability’), memorability, and selling power. As stated in the vast and extensive literature on the language of advertising (cf. Vestegaard and Schroder, 1985; Block De Behar, 1992; Ferraz Martínez, 1995; Bartha, 1997; Crompton and McAlea, 2000; Kelly-Holmes, 2005; Robles Ávila, 2005; Bathia, 2008; Cook, 2008b; Dyer, 2008; Leech, 2008; McQuarrie and Mick, 2008; Rodríguez Díaz, 2011; García Morales et al., 2016), the distinguishing traits that make it different from ordinary and regular language are the following:¹³

- (i) In opposition to everyday language, advertising language cannot be considered natural and spontaneous because it is, as any written genre, planned much in advance:

[...] language choices in advertising are not the result of a random process; they represent the attempt to use language to achieve a particular market-oriented goal, and the words present in the advertising texts are there because a very conscious decision has been taken to put them there and not to put other words there (Kelly-Holmes, 2005: 36).

- (ii) Advertising language is more connotative than denotative (cf. Bonney and Wilson, 1990: 192; Ferraz Martínez, 1995: 10-11; Kelly-Holmes, 2005: 9; Robles Ávila, 2005: 129), thus being a clear example, for Leech (2008: 174) and Dyer (2008: 299), among others, of so-called “loaded language”. As Ferraz Martínez (1995: 10) explains, it is not the products themselves, but the connotations which they evoke, defined by Cook (1992: 45) as “a vague and indeterminate world of associations quite alien to any description with pretensions to scientific rigour”, that the consumer wants to get and buy. As Ferraz Martínez (1995: 11) goes on to say, what is more important is that all advertisements are always intended to connote the same thing; basically, the excellence of the products or the services promoted, thus making them the consumer’s immediate desire. Hence, what really matters in advertisements is not their content, as Dyer (2008: 300) observes, but their linguistic style:

[...] some ads rely more on the style of language than its actual content. In ads, for, say, a foreign product like French cheese, wine or cigarettes, the speech or writing might be in the French language. We are not really expected to understand the literal meaning of the words used nor to decipher the details of the sales message but merely to recognize that it is French —a sign in itself that signifies ‘Frenchness’—.

As a consequence, “products [...] become a badge of membership”, since “by buying the product you are adopting a certain way of life” (Goatly, 2008: 92; 95). Consumers, thus, become part of a social group, “connoted”, quoting Ferraz Martínez (1995: 10), “with specific qualities such as modernity, youth, luxury, etc”. This same idea is developed by Gripsrud (2002: 82), who considers that the goods and services promoted in ads are “signs of lifestyle” and “identity”, which “tell others who we are or, rather, who we want to be”. It is in this regard that Fairclough (2008: 21) explains that advertising works ideologically by building relations, images and even the consumer.

- (iii) In order to get the aforementioned connotative effects, advertising language becomes, in Gripsrud’s (2002: 33) words, “something with which we can play and experiment”, because “it may ‘create’ reality in the sense that it can make us perceive the world differently”. Since language is used in advertising to take action and to engage with others in social life (cf. Halliday, 1970; Widdowson, 2008), “advertisement authors”, as Bathia (2008: 165) explains, “do not hesitate to sacrifice grammaticality to achieve some high-level socio- and psycholinguistic effects”. It is, nevertheless, in the varied and heterogeneous lexicon used in advertising that the creative and innovative power

of language is most clearly perceived and exploited (cf. Ferraz Martínez, 1995). Being full, on the one hand, of technical terms, and, on the other, of foreign words, which, for Kelly-Holmes (2005: 17), “give a text an elitist flavor” and for Ting-Toomey (1999: 98) are “a signal of in-group intimacy or connection”, the advertising lexicon breaks the standard norms in order to gain the potential consumers’ attention (cf. Robles Ávila, 2005: 138; Cook, 2008a: 6; Rodríguez Díaz, 2011: 173; García Morales et al., 2016: 22). In sum, as Crystal (2008: 418) argues, “language play is part of the essence of advertising”.

5. The English Language in International Advertising. The Selling Power of English

According to Bathia (2008: 166), “[t]he investigation of the language used in advertising across cultures shows that language mixing is universal and not an exceptional phenomenon”. And here, there is no doubt that English is the most favored foreign language selected for global mixing, in order to satisfy the creative and innovative needs of the advertising industry (cf. Kelly-Holmes, 2005: 182-183; Bathia, 2008: 166).

This is so because the English language, as stated in Kelly-Holmes (2005: 104; 182-183) and Bathia (2008: 166), takes a number of different, always positive, “associations” and “fetishes” —among others, modernization, Westernization, internationalism, standardization, safety, protection, domesticity, independence, informality, efficiency, competence (including scientific, technological and academic), organization, sophistication, quality, utility, physical and mental fitness, tradition, innovation and futuristic trends— which, according to Lee (2006: 59), lead to the “linguistic constructions of modernity”. Or in simpler terms, just because, as Lorenzo (1996: 11), Baumgardner (2006: 252) and García Morales et al. (2016: 19) remark, “English sells”.

Due to its selling power, therefore, it should not strike the reader as surprising that advertising, defined as “[t]he promotion of goods and services for sale through impersonal media” (Cook, 2008b: 245), makes excessive and disproportionate use of English terms. This same line of thought is brought to light by Rodríguez González (2016: 9), who openly asserts that, as expected, “los publicistas son los primeros en hacer uso desmesurado del inglés, sobre todo en el campo de la moda”.

Spanish advertising is, obviously, no exception here since it is one of the linguistic codes where most foreign words are to be found in the twenty-first century. Within them, Anglicisms occupy, undoubtedly, a privileged position for they are, as previously explained, the most numerous (cf. Gómez Capuz, 2004: 23; Rodríguez González, 2016: 7). In this regard, the report issued by the General Secretary of the Spanish Advertising Agency, Enrique Yarza (2016), in collaboration with the Complutense University of Madrid, is highly significant because it reveals that the number of English terms in Spanish advertising has remarkably increased in the last twelve years: whereas in 2003 there were around 30 brands, associated mainly with the automation, beauty and fashion sectors, using Anglicisms, in 2015 the number has risen to 322 firms, including new fields such as leisure, transport, finance, culture, the home, etc., which resort to the English language to promote their products and services. So, together with traditional slogans

like, for instance, Ford's "Go Further" and Skoda's "Simply clever", new ones, like Samsung's "Next is now" or "The Italian Sense of Beauty", used by Porcelanosa to promote the Scavolini brand of kitchen furniture, are attested nowadays. Moreover, Yarza (2016) observes that, apart from the prototypical channels used in advertising (radio, television and the written press), there are some others, newer and more modern, like walls, billboards and the shopwindows of any Spanish city, not to mention the Web, that are fostering an extreme and immoderate use of English to endorse goods and products. There, "Lunar Midnight Park" to advertise running shoes, "Must Have Vestidos Fiesta" and "Fresh and Tasty", among many others, can be found nowadays, thus becoming common mottos for any Spanish resident.¹⁴

Given the distinction between necessary and unnecessary Anglicisms previously accounted for, it should be noticed here that, although many of the English terms used in Spanish advertising are superfluous, they are not rejected by Spanish people, as Durán Martínez (2002: 44) observes, but, quite on the contrary, are widely permitted. According to García Morales et al. (2016: 23-27), the different reasons for the widespread use and acceptance of unnecessary Anglicisms in Spanish advertising can be summarized as follows:

- (i) In Spanish, English, as the official language of Anglo-American civilization, is considered to be the language of prestige. Because of this, as Lázaro Carreter (1987: 39) emphasizes, Anglicisms do not voluntarily "invade" us; we consciously "appeal" for them;
- (ii) For young people, the use of English implies linguistic snobbery (cf. Durán Martínez, 2002: 179);
- (iii) The musicality of English, together with the huge number of monosyllabic words it possesses, helps satisfy two of the basic properties of advertising language: its memorability (cf. Leech, 2008: 116) and its economy (cf. Rodríguez Díaz, 2011: 176).

In sum, I conclude, in complete agreement with Rodríguez González (2016: 12), that Anglicisms are so frequently used in Spanish advertising due to their noticeable expressive linguistic function:¹⁵

Puede decirse entonces que el empleo del Anglicismo es susceptible de adquirir una "función argótica", marcadamente expresiva. En el campo de la publicidad el Anglicismo, como en el argot, es doblemente expresivo pues, además de servir como rasgo de estilo para introducir variación, como cualidad estilística, cumple funciones pragmáticas ligadas a la intencionalidad del mensaje. Por sí mismo connota modernidad pero ayuda también a resaltar un concepto y sorprender y captar la atención del lector u oyente.

6. A Diachronic Corpus-Based Study of Present-Day Spanish Fashion Lexicon. Methodology, Data and Results

As Bathia (2008: 159) observes, Spanish and Italian are two Romance languages placed “somewhere in the middle of the continuum of closed and open languages” which “are also receptive to English mixing”.¹⁶ This statement is, in fact, verified in Gómez Cápuz (2004: 23-24), where Anglicisms are indeed highly attested in many and varied semantic fields in contemporary Spanish; namely, (i) gastronomy; (ii) fashion, where items of clothing and cosmetics are included; (iii) leisure activities, places and gambling; (iv) professions typical of Anglo-American culture; (v) politics and economics; (vi) music and urban tribes; (vii) cinema, television and shows; (viii) health and fitness; (ix) housing; (x) the military field; (xi) means of transport and urban planning; (xii) mass media and communications; (xiii) the natural and physical sciences; (xiv) and technology.

Bathia (2008: 146-147), however, argues that English is “restricted to those discourse domains which are traditionally held by English”, like, for instance, science and technology, and consequently, that “English mixing is unlikely to make its way into those domains which are held by absorbing languages for a long time”. Therefore, for him, some of the aforementioned semantic fields —romance, sophistication and fashion, in particular— are resistant to English borrowings because they have been traditionally associated with French.

Although “the hegemony of English in the fastest-growing areas of technological development”, witnessed in an incessant terminological creation in this language, as Cronin (2003: 121; 146) remarks, does not leave room for any doubt, Bathia’s (2008) assumption is not completely right because, as the following text, extracted from an Internet blog published in 2013, shows, the contemporary Spanish fashion lexicon is full of English terms:

Esta temporada las *trendsetters* han dictado sentencia: se llevan los *prints tye dye*, combinados con unos buenos tacones *peep toe*, con *blazer* y *clutch* a juego. ¡El paraíso de todas las *fashionistas*!

Si tras leer esta frase te has quedado igual que si hubieras consultado un manual de chino mandarín, tenemos la solución para ti. Últimamente hemos adoptado tantos Anglicismos, que comprender una revista de moda puede convertirse en una misión imposible, principalmente para las lectoras de a pie, que no tienen por qué conocer al dedillo la jerga del sector. (<http://www.enfemenino.com/tendencias/fashion-victim-diccionario-de-moda-d51442.html>) (Last access: April 2017)

This premise is also supported by Balteiro (2014: 160), for whom “sometimes the impression” one gets “when looking at the front page of a fashion magazine in Spanish [...] is that it is written in English, that it is a bilingual publication, or that it is mixing up both codes”.

What, for the purposes of this study, deserves special attention in the previous quotation is that the English fashion vocabulary used in it could have been replaced by Spanish original terms or expressions without great differences in meaning—a fact that

seems to suggest, thus, that the prediction concerning the incessant entrance of unnecessary Anglicisms in present-day Spanish, initially formulated by Lorenzo (1996: 18), holds true. A *trendsetter* is, in fact, an ‘iniciador o pionero de la moda’, the *prints dye* above mentioned are simply ‘estampados’, the *tacones peep toe* refer to those ‘zapatos con la puntera abierta’, the traditional *blazer* denotes a ‘chaqueta’, the *clutch* alludes to a ‘bolso de mano’, and finally, *fashionistas* points to ‘gente adicta a la moda’.

6.1. Data and Methodology

In order to verify Lorenzo’s (1996: 18) hypothesis, I will analyze the frequency of occurrence that the following 14 superfluous “patent Anglicisms” (cf. Pratt, 1980; Lorenzo, 1996) have at different stages in the evolution of Peninsular Spanish: *blazer*, *celebrity*, *clutch*, *cool*, *fashion* and its derivatives *fashionable* and *fashionista*, *jeans*, *nude*, *photocall*, *shorts*, *sporty*, *trench* and *trendy*. With the exception of *fashionista*, which adds the Spanish suffix *-ista* to an English stem, all the others maintain unaltered in Spanish their English form, thus belonging to the group of Lorenzo’s (1996) “raw Anglicisms” (*Anglicismos crudos*) and to that of “foreign words” (*extranjerismos*) in Alfaro’s (1948) classification. As explained earlier, they are gratuitous Anglicisms because they have equivalent original Spanish terms or expressions to name the extra-linguistic reality they designate: ‘chaqueta (deportiva, de colegio)’, ‘famoso’, ‘bolso de mano’, ‘de moda’, ‘moda’, ‘moderno’, ‘persona adicta a la moda’, ‘vaqueros’, ‘color carne’, ‘sesión de fotos’, ‘pantalones cortos’, ‘deportivo’, ‘trinchera, gabardina’ and ‘moderno’.

Their selection, instead of that of others, is mainly motivated by their recurrent presence in Spanish shopwindows, websites and current fashion magazines, as illustrated in the following examples retrieved at one single mouse click from the *Hola Fashion* magazine website:

- (1) Cómo ser **‘cool’** con un uniforme deportivo. Te contamos como una biker de cuero o unas ‘sneakers’ de aire motero pueden transformar un ‘look’ **‘sporty’** desenfadado en sofisticado. (13/01/2017)
- (2) El **‘blazer’**: Cómo lucirlo [...] Combínala con básicos de estética casual. Camiseta+**jeans**+sneakers, una combinación perfecta para restar seriedad a la prenda y conseguir un aire **cool**. (20/10/2016)
- (3) ¿Qué **‘clutch’** se ha convertido en la ‘pieza’ imprescindible de todos los ‘looks’? Este diseño, inspirado en las piezas de Lego, es uno de los favoritos de las **‘celebrities’**. (29/10/2013)
- (4) 100% **‘trendy’**, de la mañana a la noche. **Shorts** y sandalias planas para el día, cuñas altas para unos refrigerios afterwork, tacones y toques de brillo para un look de noche. (13/05/2014)
- (5) Esta temporada los estilos colegial y deportivo se han convertido en los hits de todas las colecciones: el estampado de cuadros y las sudaderas invaden los armarios de las **fashionistas y celebrities**. (18/11/2013)

- (6) **Photocalls**, alfombras rojas, front-rows... Espacios que se convierten en principales testigos de qué prendas y complementos llevan a la reina de la armada **fashion** neoyorquina a reutilizar sus prendas fetiche. (18/11/2014)
- (7) [...] el **trench** fue una de las piezas clave. [...] la diseñadora [...] le ha dado a esta prenda un toque más actual [...] aunque conservando su característico color **nude**. (11/09/2017)

Though the lexical inventory under study is clearly unbalanced in terms of the lexical category of its components, its disproportion is justified since it confirms the ‘hierarchy of borrowability’, described, among others, in Pulcini et al. (2012: 9) and MacKenzie (2012: 31), where foreign concrete nouns come first, closed-class words at the end, and adjectives, together with abstract nouns, verbs and adverbs, in the middle. Notice that it consists, specifically, of nine nouns, five of which denote different items of clothing — *blazer*, *clutch*, *jeans*, *shorts*, and *trench*— and four ones also highly frequent in the contemporary fashion world —*fashion*, *fashionista*, *celebrity*, and *photocall*—, along with five adjectives that, by qualifying different fashion styles —*cool*, *fashionable*, *sporty*, and *trendy*— or describing a particular color —*nude*— are, likewise, visible representatives of this particular lexical field. *Cool* and *trendy*, in particular, are included in the group of adjectives which MacKenzie (2012: 38) describes as “widely diffused in Scandinavia”, as well as in some of the Spanish lexicographical sources examined, as will be detailed later.

As most of the chosen terms are nouns, they have been looked for in the corpora used in my analysis in the singular and plural. This double search has also been applied to the adjectival Anglicisms in my inventory because, although in Spanish they should be invariable in number, following their English morphological behaviour, their inflected pluralization is sometimes a common well-founded fact (cf. Rodríguez González, 2017: 327-328). Therefore, since some Anglicisms in the plural “give rise, especially in writing, to noticeable variation”, as Rodríguez González (2017: 299-300) highlights, all the possible written plural forms in which the terms in my list can appear —both accepted and non-accepted in standard Peninsular Spanish by the Royal Academy— have been considered. In particular,

- (i) For the four Anglicisms whose singular form ends in a consonant different from *-ch* —*blazer*, *fashion*, *cool*, and *photocall*— two different plural orthographic forms have been examined: (i) the supposed accepted form which, after the English canonical rule of pluralization, adds an *-s* to this kind of words; (ii) and the non-accepted version ending in *-es*, wrongly applied to them by analogy with the Spanish plurals of the words that, like, for instance, *color*, *camión* and *farol*, finish in the same consonants as the foreign terms.
- (ii) Two plural forms have also been searched for the two nouns —*clutch* and *trench*— whose singular form ends in the affricate sound [tʃ], graphically represented with the spelling *-ch*: (i) their approved plural form in *-es*, also borrowed from English; (ii) and their wrong plural form in *-s*, which results from overgeneralizing the Spanish pluralization rule applied to those English terms, such as *pub* and *club*, among others, whose singular form ends in the plosive [b] sound.

- (iii) Although the plural of *jeans* and *shorts* is strongly consolidated in Spanish and thus far from being a problem, owing to the inexistence of their singular form in English, these two Anglicisms have also been analyzed in the singular because their Spanish counterparts do show number contrast: ‘pantalón(es) vaquero(s)/ corto(s)’.
- (iv) Since the pluralization of those Anglicisms that finish in a vowel different from *-y* —*fashionable*, *fashionista* and *nude*— is not problematic, for them only one single form, the one that ends in *-s*, has been taken into consideration. However, for the three Anglicisms with a final *-y* vocalic sound —*celebrity*, *sporty*, and *trendy*—, four plural orthographic forms have been examined: (i) the original English plural in *-ies*; (ii) their accepted Spanish plural form in *-is*; (iii) and their two wrong plural versions which end, in turn, either in *-ys* or *-yes*.

In addition, the corpora where their real use and productivity have been looked for are the following ones: (i) the *Corpus del Español*, created by Professor Davies (2002), which contains more than 100 million written and spoken words compiled from different varieties of Spanish (approximately, 50% Peninsular Spanish and 50% Latin American Spanish), dating from the thirteenth up to the twentieth centuries; (ii) the *CORDE* corpus, which consists of 250 million Peninsular (74%) and Latin American (26%) Spanish forms extracted from texts of very different genres written during the period that goes from the very beginning of the language, dated in 950 in Oncins-Martínez (2012: 20), to 1974;¹⁷ (iii) the annotated version of the *CREA* corpus, first published in November 2015, with more than 126 million written words of Peninsular and Latin American Spanish (each variety accounting for 50%) produced in the period that goes from 1975 to 2000;¹⁸ (iv) and finally, the *CORPES XXI* corpus, which comprises 225 million written and oral forms from Peninsular (30%) and Latin American (70%) Spanish belonging to the period ranging from 2001 to 2015:

	Time span	Number of words	Number of words in Peninsular Spanish	Registers
<i>Corpus del Español</i>	13th c.-20th c.	+100 million words	+50 million words	written/spoken
<i>CORDE</i>	950-1974	250 million words	185 million words	written
<i>CREA (annotated)</i>	1975-2000	+126 million words	63 million words	written
<i>CORPES XXI</i>	2001-2015	225 million words	67,5 million words	written/spoken

Table 1: Periods, registers and number of words in the corpora

Although these four corpora provide an online interface, they are somehow different. For Oncins-Martínez (2012: 220), for instance, the one in the *CORDE* corpus “is not as user-friendly as those in other similar —if smaller— corpora (e.g. Davies’s *Corpus del*

Español)” because in it relative frequencies must be “manually calculated using the figures given in the statistics section”. In the three other corpora, however, both the raw and normalized frequencies of the results obtained are offered. Another significant difference among these corpora concerns the search options available in them: time, types of texts and registers and varieties of Spanish. Specifically, this last variable, essential for the purposes of my work, is present in the three Spanish Royal Academy corpora, but absent from the *Corpus del Español*. Therefore, with this in mind, each of the attestations retrieved for each of the Anglicisms under study has been closely analyzed so as to exclude the following occurrences:

(i) Those instances in which the English term is a proper name:

- (8) Por el **Portland Trail Blazers** volvió a destacar su bombardero blanco Kiki Vandeweghe [...]. (*CREA*, Prensa, 1986)
- (9) **Jean Chalon** estalla con la risa jovial, inofensiva y contagiosa del chico mimado por la fortuna que mantiene intactas sus ambiciones juveniles. (*Corpus del Español*, Oral: interview *ABC*)

(ii) Those examples where the word examined appears in a context that is not Spanish:

- (10) Deinceps per praedicationem qualis dicta est, populus qui est pronissimus ad suggestiones, curiositates et sortilegia credenda, reditur proclivior dum talia palam accipit in sermonibus, in quibus fides recta solido et **nude** praedicanda foret, et in ea populus nutriendus plus, quam in fabulis ancilibus. (*CORDE*, 1818)

(iii) Those ones in which the meaning of the Anglicism is not the one in which we are interested here, like *cool*, for instance, referring to jazz music in (11), or *clutch* in (12), denoting, in turn, ‘embrague’:

- (11) En la actualidad todas las corrientes conviven y evolucionan en sus estilos (ragtime, blues, dixieland, big band, espirituales, bebop, **cool** jazz, freejazz, electric jazz, etc.) bajo el amplio paraguas del término jazz, [...]. (*CREA*, Miscelánea, 1998)
- (12) El primero [...] está formado por el volante del motor, un plato conductor que gira junto a éste y un disco conducido o de **clutch** situado entre ambos que está unido al eje primario o flecha de mando de la caja de cambios (*Corpus del Español*, http://es.encarta.msn.com/artcenter_/browse.html)

(iv) And finally, also those instances from the *Corpus del Español* in which the English borrowings have been attested in a variety different from Peninsular Spanish:

- (13) Sepp Blatter se presentó en punto de las 10:30 de la mañana en un elegante **blazer** beige. (*Corpus del Español*, Mex: Yucatán: 97Jun28)
- (14) Visten por una moda importada; sus blue-**jeans** (bluyín) y sus florecitas son importadas de... Estados Unidos, de Miami o de Nueva York; (*Corpus del Español*. Habla Culta: Caracas: M1)

<i>Sporties</i>	-	-	-	-	-	-	-	-
<i>Sportis</i>	-	-	-	-	-	-	-	-
<i>Sportys</i>	-	-	-	-	-	-	-	-
<i>Sportyes</i>	-	-	-	-	-	-	-	-
<i>Trench</i> (1917)	-	-	-	-	1	0.01	4	0.05
<i>Trenches</i>	-	-	-	-	-	-	-	-
<i>Trenchs</i>	-	-	-	-	-	-	1	0.01
<i>Trendy</i> (1962)	-	-	-	-	1	0.01	24	0.35
<i>Trendies</i>	-	-	-	-	-	-	1	0.01
<i>Trendis</i>	-	-	.	-	-	-	-	-
<i>Trendys</i>	-	-	-	-	-	-	1	0.01
<i>Trendyes</i>	-	-	.	-	-	-	-	-
TOTAL	16	0.32	28	0.141	95	1.46	566	8.25

Table 2: Raw and normalized frequencies of the Anglicisms studied

This initial result is, obviously, the expected one for *clutch*, *fashionista*, *nude*, *photocall*, *trench*, and *trendy* since, having entered English along the twentieth century, as indicated in the *OED* (2017) and illustrated between brackets next to them in Table 2, their presence in Peninsular Spanish can only be verified in the contemporary period. To my mind, this is also probably the reason why, with the exception of *trendy*—only registered in Rodríguez González and Lillo Buades (1997)—none of the others has a place in the bibliography reviewed.

Fashionista, specifically, is a particular case to consider because, though a clear Anglicism coming from *fashionist* (an English term first attested probably in ?1624), it has become part of the English lexicon in 1992 as a “borrowing coming from Spanish” (*OED*, 2017). It is, thus, its recent entry date in English that explains why its use in Peninsular Spanish, (15-16), is only registered in the *OLD* (2017):

- (15) [...] éste fue publicado en Inglaterra bajo el transparente alias hembra de Danuta de Rhodes: una precoz **fashionista** que ya a sus doce años firmaba artículos en las mejores revistas de moda, [...] (*CORPES XXI*, Prensa, 2005)
- (16) Puede, pero dudo que las **fashionistas** vayan a agotar las existencias de esta colección algo ochentera y más bien sosa. (*CORPES XXI*, Internet, 2006)

What surprises, consequently, about *fashionista* is the attestation dates of the 11 singular and plural instances found with it, which, as shown in Table 3, occur much later in time in Peninsular Spanish, over a recent short span of eight years (2005-2012), than the Spanish borrowing in English:

	<i>Corpus del Español</i> (13th c.-20th c.)	<i>CORDE</i> (950-1974)	<i>Annotated CREA</i> (1975-2000)	<i>CORPES XXI</i> (2001-2015)
<i>Fashionista</i> (1992)	-	-	-	7 Books (1 instance) 2006 Press (4 instances) 2005 (2), 2009, 2010 Internet (2 instances) 2012 (2)
<i>Fashionistas</i>	-	-	-	4 Books (1 instance) 2008 Press (1 instance) 2008 Internet (2 instances) 2006, 2012

Table 3: Corpora distribution of *fashionista(s)* per year/register

The closest Anglicisms in time to *fashionista* of the ones in my inventory are *photocall* and *trendy*, which, according to the *OED* (2017), have entered English approximately thirty years earlier, after the mid-twentieth century; in 1958 and 1962, respectively. Despite a temporal interval of only five years between them, there are significant differences between the distribution and number of attestations obtained for each of these Anglicisms:

	<i>Corpus del Español</i> (13th c.-20th c.)	<i>CORDE</i> (950-1974)	<i>Annotated CREA</i> (1975-2000)	<i>CORPES XXI</i> (2001-2015)
<i>Photocall</i> (1958)	-	-	-	10 Books (2 instances) 2012 (2) Press (5 instances) 2009, 2012 (4) Internet (3 instances) 2011, 2012 (2)
<i>Photocalls</i>	-	-	-	1 Internet (1 instance) 2012
<i>Trendy</i> (1962)	-	-	1 Book (1 instance) 1997	24 Books (9 instances) 2008 (2), 2010, 2012 (5), 2013

				Press (12 instances) 2002, 2003 (2), 2006 (3), 2008, 2009 (2), 2010 (2) 2012 Internet (3 instances) 2011 (2), 2012
<i>Trendys</i>	-	-	-	1 Internet (1 instance) 2010
<i>Trendies</i>	-	-	-	1 Press (1 instance) 2002

Table 4: Corpora distribution of *photocall* and *trendy* per year/register

Though documented earlier in English than *trendy*, *photocall* in my corpus search has yielded fewer instances: 11 versus 27. All of them are, furthermore, quite new and recent, being found, except for example (17), in the second decade of the twenty-first century:

- (17) Unas fotografías de la pareja paseando por la noche han puesto de nuevo a la canaria en primera línea del rostro más buscado para los ‘**photocall**’, [...]. (*CORPES XXI*, Prensa, 2009).

In my view, there are two reasons that account for the uneven results obtained for *photocall* and *trendy*: on the one hand, their different lexical categories; and on the other, the number of varying plural forms they may have. As a concrete noun, *photocall* just denotes a discrete entity —namely, “An occasion arranged to enable a celebrity or celebrities to pose for photographers” (*OED*, 2017)— and, consequently, has a very restricted usage. Conversely, as an adjective that means “Fashionable, up to date, following the latest trend” (*OED*, 2017), *trendy* can modify a wide range of entities, as seen in (18-20), thus multiplying its possibilities of use:

- (18) En los albores del año 2000 se ha introducido el gusto por las opciones múltiples. No es posible siquiera ser **trendy**; adictos a la última moda; (*CREA*, Libro, 1997)
 (19) En resumen, que esta temporada, los chicos podéis estar cómodos y abrigaditos a la par que **trendys**, [...]. (*CORPES XXI*, Internet, 2010)
 (20) Sebastián Pila es el encargado de Doble A, una de las tiendas de moda mas **trendies** de Madrid. (*CORPES XXI*, Prensa, 2002)

Though both Anglicisms have been described before as problematic in terms of their pluralization in Spanish, not all of their possible plural forms have been attested in my study. As seen in the previous examples, for *trendy* only its English plural version, (20), and the unaccepted Spanish form in *-ys*, (19), out of the three potential ones, have been

documented. As regards *photocall*, of its two probable alternatives, only its accepted plural form in standard Spanish has been obtained, (21):

- (21) [...] en los «saldillos de prensa» se venden las prendas que se han utilizado durante tres o cuatro meses en producciones de moda de revistas, o para vestir a famosos en los «**photocalls**». (*CORPES XXI*, Internet, 2012)

Less numerous are the instances obtained for *clutch* and *trench*, despite being older in English than the terms previously discussed. Specifically, nine attestations have been found with the former and six with the latter. With only one exception for *trench*, they all are located in the twenty-first century:

	<i>Corpus del Español</i> (13th c.-20th c.)	<i>CORDE</i> (1950-1974)	<i>Annotated CREA</i> (1975-2000)	<i>CORPES XXI</i> (2001-2015)
<i>Clutch</i> (1950)	-	-	-	8 Books (1 instance) 2009 Internet (7 instances) 2011 (5), 2012 (2)
<i>Clutches</i>	-	-	-	1 Internet (1 instance) 2010
<i>Trench</i> (1917)	-	-	1 Book (1 instance) 1997	4 Books (3 instances) 2007, 2008, 2009 Internet (1 instance) 2010
<i>Trenchs</i>	-	-	-	1 Press (1 instance) 2006

Table 5: Corpora distribution of *clutch* and *trench* per year/register

What deserves special attention in relation to both Anglicisms is that they have entered Pensinsular Spanish as the shortened versions of *clutch bag* and *trench coat*, as seen in (22-23), which, curiously enough, however, are not registered as such in the English-Spanish bilingual lexicographical sources examined. In the *OED* (2017), in turn, their presence is confirmed but in reference to the aforementioned compound terms:

- (22) El **clutch** de volantes de Yves Saint Laurent. ¡Es París! (*CORPES XXI*, Libro, 2009)
 (23) Llevaba un **trench** de color verde jade que se quitó para entrar en el vehículo [...]. (*CORPES XXI*, Libro, 2008)

On the other hand, their plural forms, though just attested once with each Anglicism, are not homogeneous since, whereas the one found with *clutch* adopts the approved version in Spanish of the English plural, (24), the one with *trench* wrongly overgeneralizes the Spanish rule for some foreign plurals, thus becoming the non-accepted standard alternative, (25):

- (24) En su catálogo podemos encontrar shopping bags enormes, [...], o bandoleras en charol de lo más chic, pasando por algunas piezas más clásicas pero atemporales, como los **clutches** con forma de sobre en ante negro. (*CORPES XXI*, Internet, 2010)
- (25) Bombers revisados, faldas acolchadas y una gran variación de **trenchs** en un desfile en el que el tejido de gabardina fue muy especial [...] (*CORPES XXI*, Prensa, 2006)

The meaning of *nude* that refers to a particular colour is dated in English in 1922. However, it is not until 2003, as illustrated in Table 6, that it is attested in my corpus search, and this accounts for its complete absence in all the dictionaries consulted:

	<i>Corpus del Español</i> (13th c-20th c.)	<i>CORDE</i> (950-1974)	<i>Annotated CREA</i> (1975-2000)	<i>CORPES XXI</i> (2001-2015)
<i>Nude</i> (1922)	-	-	-	15 Books (1 instance) 2011 Press (3 instances) 2003, 2009, 2011 Internet (11 instances) 2010, 2011 (10)

Table 6: Corpora distribution of *nude* per year/register

It is interesting to highlight here that *nude* has been found with exactly the same form in the singular, (26), and plural, (27), thus following the morphological invariability in terms of number of English adjectives. No single instance of its Spanish plural version in *-s* has been attested:

- (26) Combina después dos tonos de fondo: "El 030 sand [...]. Y el 020 **nude**, con el que se obtienen destellos luminosos en zonas claves del rostro: [...]. (*CORPES XXI*, Prensa, 2003)
- (27) [...] este modelo en color coral es perfecto para los looks bohemian chic. Combínalo con tonos **nude** e irás perfecta. (*CORPES XXI*, Internet, 2011)

In opposition to the former ones, almost all the Anglicisms in my inventory whose entry date in English is registered at some time in the nineteenth century—in chronological order, *shorts* (1826), *celebrity* (1831), *jeans* (1843), *blazer* (1880), *cool* (1884), and *sparty* (1895)—are commonly present in the bibliographical references on the topic, *celebrity* and *sparty* being the only exceptions in this regard.

The non-attestation of *sparty* in the bibliography on Spanish Anglicisms, despite the presence in it of its morphological root —*sport*— (cf. CLAVE, 1996; Lorenzo, 1996; Rodríguez González and Lillo Buades, 1997; DNLE, 1998; Gómez Capuz, 2004; DUE, 2007; DRAE, 2014), matches the results derived from my corpus-based study, where there is not a single trace neither of its singular nor of its three possible plural forms. The reason that, in my view, accounts for this particular fact is to be found in Lorenzo (1996: 416-418), who, drawing on the research on *sport* and its derivatives carried out by Fernández García (1972), concludes that “el adjetivo correspondiente, *sportivo*, *esportivo*, adaptado del inglés *sportive*, aparece en 1895, y está bien documentado hasta ir cediendo el puesto, desde 1918, a *deportivo*”. This same explanation could be extended to *sparty* since in contemporary Spanish *deportivo*, coming from *deporte* —an English calque from *sport* completely adapted to Spanish orthographic norms— has exactly the same meaning as the English term *sport* —this term being used, as stated in the DRAE (2014), as an adjective to refer to casual items of clothing. The complete absence of *sparty* in my corpus search seems, therefore, to suggest that if Spanish possesses an English term adapted to its spelling rules, this alternative will be preferred over its equivalent raw Anglicism.

With *celebrity*, however, something different occurs since 55 singular and plural instances of this particular Anglicism have been attested along the first fifteen years of the twenty-first century. Given that its earliest occurrence dates, as shown in Table 7, from 2006, its complete non-existence in the Spanish lexicographical sources consulted is justified:

	<i>Corpus del Español</i> (13th c.-20th c.)	<i>CORDE</i> (950-1974)	<i>Annotated CREA</i> (1975-2000)	<i>CORPES XXI</i> (2001-2015)
<i>Celebrity</i> (1831)	-	-	-	9 Books (1 instance) 2009 Press (5 instances) 2008, 2009, 2010, 2012 (2) Internet (3 instances) 2010, 2011, 2012
<i>Celebritis</i>	-	-	-	1 Press (1 instance) 2012
<i>Celebrities</i>	-	-	-	42 Books (12 instances) 2006 (5), 2007 (3), 2010 (2), 2011 (2) Press (12 instances) 2007 (2), 2008 (2), 2010 (3), 2011 (4), 2012 Internet (18 instances) 2006 (2), 2010 (3), 2011 (11), 2012 (2)

Table 7: Corpora distribution of *celebrity* per year/register

Moreover, *celebrity* stands out from the Spanish Anglicisms under study, together with *jeans*, in two significant ways: first, because it is the only one whose first attestation in Peninsular Spanish is found in the plural; and secondly, because more plural than singular forms have been obtained with it: specifically, 43 versus nine. As illustrated in (28-29), its plural presents two of the three possible orthographic forms it could have *a priori*: the English canonical plural in *-ies*, being the most recurrent, and the approved Spanish plural in *-is*:

- (28) [...] lo que a mí más me gustó, son los vestidos de noche. Eso y los botines-mocasines. En fin, de momento ha convencido a las **celebrities**, pero habrá que ver cómo funciona. (*CORPES XXI*, Internet, 2006)
- (29) [...] estará dedicada a las noticias más ligeras, vinculadas a aspectos sociales y relacionadas con las **celebritis**. (*CORPES XXI*, Prensa, 2012)

As regards *cool*, its only presence in Rodríguez González and Lillo Buades (1997) and *DUE* (2007) has to be considered somewhat surprising for two reasons: first, because its first attestation in Peninsular Spanish dates from 1980; and secondly, because it is the most recurrent Anglicism of the ones examined, with 186 occurrences along the last quarter of the twentieth century and the first thirteen years of the twenty-first century:

	<i>Corpus del Español</i> (13th c.-20th c.)	<i>CORDE</i> (1950-1974)	<i>Annotated CREA</i> (1975-2000)	<i>CORPES XXI</i> (2001-2015)
<i>Cool</i> (1884)	-	-	<p>12</p> <p>Libros (4 instances) 1994, 1997, 1998, 2000</p> <p>Press (8 instances) 1980 (2), 1981, 1994, 1995 (4)</p>	<p>174</p> <p>Books (50 instances) 2001 (3), 2006 (5) 2007, 2008 (6) 2009 (11), 2010 (2) 2011 (13), 2012 (8) 2013</p> <p>Press (93 instances) 2002 (2), 2003(8) 2004 (7), 2005 (16) 2006 (4), 2007 (6) 2008(12), 2009 (19) 2010 (6), 2011 (4) 2012 (9)</p> <p>Internet (31 instances) 2002, 2005 (3) 2006 (2), 2008 2009, 2010 (4) 2011 (15), 2012 (4)</p>

Table 8: Corpora distribution of *cool* per year/register

To my mind, its increasingly extended use in contemporary Peninsular Spanish is to be accounted for on the basis of the following arguments: (i) first, its morphological invariability in terms of number to cover, as any English adjective, both its singular and plural uses, (30-31); (ii) secondly, its familiar and colloquial tone, frequently highlighted in the bilingual lexicographical sources consulted, which helps foster its diffusion; (iii) and finally, its many positive connotations which, encoded in Spanish by means of different words (*fabuloso*, *chachi*, *guay*, *de aupa*, *de moda*, among others), make its use apt in a wide array of contexts to refer both to personal, (30) and nonpersonal, (31-33), referents, among which those related to the fashion lexicon are included, (33):

- (30) Salen del salón de Valentino, llevan abrigos de cachemir espectaculares y sus tarjetas de crédito tintinean. Pero no, no son «cool». (*CORPES XXI*, Prensa, 2005)
- (31) Es una bebida extravagante [...], que tiene su particular reino en Le Bar du Plaza Athennée, el bar más **cool** de París, [...], en el que se puede ver a Karl Lagerfeld tomando una copa con su amiga Carolina de Mónaco. (*CORPES XXI*, Libro, 2006)
- (32) Registrar dominios a nombre de un hijo recién nacido o de la pareja sentimental es el obsequio más **cool** en Estados Unidos. (*CORPES XXI*, Prensa, 2007)

- (33) [...] una sombrerería [...] en donde podréis encontrar sombreros tipo Borsalino, Boinas, Mayser, Stetson o Kangol. Este complemento te dará un look **cool** y debe de formar parte del estilismo veraniego. (*CORPES XXI*, Internet, 2010)

Despite being as superfluous as the other Anglicisms having entered English in the nineteenth century, *jean(s)*, *short(s)* and *blazer* are usually included, in opposition to these others, in most of the sources examined (cf. *CLAVE*, 1996; *DUE*, 1996; Lorenzo, 1996; Rodríguez González y Lillo Buade, 1997; *DNLE*, 1998; Gómez Capuz, 2004; *DRAE*, 2014; Yarza, 2016). Though their presence in present-day Peninsular Spanish can be considered somewhat alike in terms of quantity, as will be immediately shown, only *jeans* and *shorts* have been attested in the two corpora that cover the earlier periods of this Spanish variety, though with an extremely low frequency of occurrence. They are, furthermore, together with *fashion*, the only Anglicisms in my inventory which have been found in the oral registers of the language, (34-35):

- (34) [...] en ese momento llevaba un maletín con un cuchillo de cocina, un **short** de niña de color rosa con dos agujeros, que se ponía como mascara [...] (*Corpus del Español*, Oral: CNOT034A)
- (35) El escritor está vestido en su estilo casi juvenil: una gastada camisa roja, un buzo borravino, «**jeans**» y zapatones. (*Corpus del Español*, Oral, interview *ABC*)

As regards the first of the terms, *jeans*, Pratt (1980: 215) highlights its unpopularity and its consequent scarce use in Peninsular Spanish, which Lorenzo (1996: 256) attributes to the widespread usage of the general Spanish nouns *tejanos* y *vaqueros*. Though Pratt's (1980: 215) assumption holds true until the time of his publication, after that date, as illustrated in Table 9, the diffusion of *jean(s)* has undergone a noticeable growth; whereas up to the year 1980 only six attestations have been found, 77 instances have been retrieved for the period comprised between 1981 and 2012:

	<i>Corpus del Español</i> (13th c.-20th c.)	<i>CORDE</i> (1950-1974)	<i>Annotated CREA</i> (1975-2000)	<i>CORPES XXI</i> (2001-2015)
<i>Jean</i>	-	-	1 Press (1 instance) 1998	-
<i>Jeans</i> (1843)	2 Oral (2 instances)	3 Fiction (3 instances) 1966 (3)	28 Books (11 instances) 1981 (2), 1982 (2) 1989, 1991, 1992 (2), 1997 1998, 2000 Press (17 instances) 1980, 1986 1994 (4), 1995(3) 1996 (2), 1997(4) 1998 (2)	49 Books (25 instances) 2001 (12), 2004 2005, 2006, 2008 2009, 2010, 2011 2012 Press (7 instances) 2004, 2005, 2006 2007, 2008, 2010 Internet (16 instances) 2010, 2011 (15) Oral (1 instance) 2003

Table 9: Corpora distribution of *jean(s)* per year/register

In relation to *short(s)*, Lorenzo (1996: 396) remarks that its singular and plural forms have been registered apart, as two different entries, in Fernández García (1972); the former in 1931 to designate a ‘short film’ and the latter in 1934 to refer to ‘short trousers’. Though the occurrences of its plural form, except for the one retrieved from the *Corpus del Español*, whose exact date is not known, are indeed posterior to the time indicated (the first ones date from 1966), the singular attestations of this Anglicism, referring to ‘pantalón corto’, (36), reveal that their entrance in Peninsular Spanish is earlier in time (1951) than that of the plural:

	<i>Corpus del Español</i> (13th c.-20th c.)	<i>CORDE</i> (1950-1974)	<i>Annotated CREA</i> (1975-2000)	<i>CORPES XXI</i> (2001-2015)
<i>Short</i>	1 Oral (1 instance)	2 Fiction (2 instances) 1951, 1953	14 Books (13 instance) 1976, 1989, 1991 1993, 1995 (9) Press (1 instance) 1996	3 Internet (3 instances) 2003 (2) 2004
<i>Shorts</i>	1 Fiction (1 instance)	5 Fiction (5 instances) 1966 (3) 1972 (2)	22 Books (11 instances) 1977, 1978 (2) 1981 (2), 1990 (2) 1991, 1999 (2) 2000 Press (11 instances) 1989 (2), 1990 1994, 1995 (2) 1996 (2), 1997 (3)	35 Books (11 instances) 2001 (2), 2004 2005 (2), 2006 2009, 2011 2012 (3) Press (10 instances) 2002, 2003 (3) 2005, 2006 2007, 2010 (2) 2012 Internet (14 instances) 2008, 2010 (4) 2011 (7) 2012 (2)

Table 10: Corpora distribution of *short(s)* per year/register

On the other hand, though both Anglicisms appear surprisingly in the singular in the *DRAE* (2014), if with a note that explains that they mean the same as in the plural, their plural form is the most recurrent one in Peninsular Spanish nowadays. With *short(s)*, in particular, 20 singular and 83 plural instances have been found; furthermore, excluding the tokens retrieved from the *Corpus del Español*, whose exact attestation dates are not known, most of them have been documented in the period that goes from 1975 to our days (2015); specifically, 17 singular and 77 plural attestations. The two remaining singular forms, along with the five other plural ones, have been dated, in turn, at some point earlier in time. With *jean(s)*, however, the contrasts in number are even more pronounced since, against the one singular form documented in 1998, 82 plural instances have been obtained; in particular, 77 in the period that covers the last quarter of the twentieth century and the beginning of the twenty-first century and only three, the two attestations from the *Corpus del Español* excluded, before that time interval.

Finally, as regards *blazer*, it should be highlighted that its diffusion in Peninsular Spanish has undergone a noticeable rise over time; whereas it was completely absent from 950 to 1988, the year in which its first occurrence is registered in the *CREA* corpus, its presence in the period that goes from 1994 to 2012 has been widely attested, with 76 singular and plural occurrences:

	<i>Corpus del Español</i> (13th c.-20th c.)	<i>CORDE</i> (1950-1974)	<i>Annotated CREA</i> (1975-2000)	<i>CORPES XXI</i> (2001-2015)
<i>Blazer</i> (1880)	-	-	10 Books (2 instances) 1988, 1997 Press (8 instances) 1994, 1996, 1997 (4)	51 Books (42 instances) 2003, 2004, 2005 2006 (2), 2008, 2009, 2010, 2011 (33), 2012 Press (3 instances) 2006, 2010 (2) Internet (6 instances) 2011 (4) 2012 (2)
<i>Blazers</i>	-	-	6 Books (5 instances) 1997 Press (1 instance) 1997	10 Books (7 instances) 2011 Internet (3 instances) 2011

Table 11: Corpora distribution of *blazer* per year/register

The reason which, I believe, explains this quantitative contrast is the progressive loss, as seen in (36-37), of some of the connotations usually ascribed to this Anglicism in the lexicographical sources over time: among others, the blue colour and the badge typical of the blazer of school uniforms (cf. Lorenzo, 1996: 128; *DUE*, 2007), and/or its sporty style (cf. *DUE*, 2007; *DRAE*, 2014):

- (36) Entraron en la tienda de Adolfo Domínguez. [...] Gaspar le metió en el vestidor tres prendas más. Un **blazer** blanco, una rebeca de verano, una faldita corta que Julia jamás se hubiera probado. (*CORPES XXI*, Libro, 2006)
- (37) Confío en que mi vestido de Mango [...] con el **blazer** (es una americana, pero se dice así porque es fashion) de color beige con etiqueta de Stella McCartney [...] lleven aire acondicionado. (*CORPES XXI*, Libro, 2010)

Finally, the two oldest Anglicisms in my list are *fashion* and *fashionable*, which according to the *OED* (2017) have entered English, respectively, in the fifteenth (1490) and seventeenth (1630) centuries. *Fashionable*, in particular, has undergone a progressive decline over time since, against the 27 singular and plural occurrences attested in the earliest stages of Peninsular Spanish examined (20 in the nineteenth century and only seven in the twentieth century), it is entirely absent in later periods of time, given the results derived from the *CREA* and the *CORPES XXI* corpora:

	<i>Corpus del Español</i> (13th c.-20th c.)	<i>CORDE</i> (950-1974)	<i>Annotated CREA</i> (1975-2000)	<i>CORPES XXI</i> (2001-2015)
<i>Fashionable</i> (1630)	7 Fiction (7 instances) 1836 (3), 1856 1864, 1961 (2)	13 Fiction (13 instances) 1832, 1846 1847-1857 1849, 1852 1869, 1872-1878 1876, 1897 1900, 1904 1912 (2)	-	-
<i>Fashionables</i>	3 Fiction (3 instances) 1823, 1836 (2)	4 Fiction (4 instances) 1835, 1849 1885 1944-1949	-	-

Table 12: Corpora distribution of *fashionable* per year/register

My results do, therefore, confirm Lorenzo's (1996: 209) observation that this Anglicism is an old-fashioned term in contemporary Spanish and in French. My results differ, however, in the registration dates of *fashionable* provided. For Lorenzo (1996: 209), its earliest attestation as a noun is found in Larra (1835) and as an adjective in Valera (1847). In my study, in turn, an earlier nominal instance of *fashionable*, and an adjectival one, have been found; as illustrated in (38-39), the former in Mesonero Romanos (1832) and the latter in Navarro Villoslada (1846):¹⁹

(38) y sin haber salido de Zaragoza, afectaba ya los usos de un **fashionable** de Londres [...] (*CORDE*, 1832)

(39) Hemos advertido ya la mezquina rivalidad **fashionable** que reinaba en punto a trajes entre españoles y franceses [...] (*CORDE*, 1846)

There are two clear reasons that, in my opinion, may explain the decrease in the use of *fashionable* in present-day Peninsular Spanish: first, the frequent adjectival use, illustrated in (40-41), which the English noun *fashion* has developed in this Spanish variety throughout time, attested, in fact, in 73 occurrences out of the 113 singular instances obtained with *fashion*:

(40) [...] atracción por el rap, ese look negrata que viene ahora, que será trendy (de hecho Madonna va a sacar un disco de hiphop) y me apetece ser un poco visionaria. Eso sí lo veo **muy fashion**. (*CORPES XXI*, Press, 2003).

(41) [...] la historia de un chico normal, de un robot bastante chatarrero y **poco fashion** que, gracias a su ingenio y a su tesón, logrará la fama, [...] (*CORPES XXI*, Press, 2005).

And secondly, the rising occurrence of *fashionista* over time, previously described, as well as that of *fashion*. As shown in Table 13, against the three singular occurrences of

fashion found in the *Corpus del Español* and the *CORDE* corpus (two in the nineteenth century and one in the twentieth century), 115 singular and plural instances have been attested along the first thirteen years of the twenty-first century covered in the *CORPES XXI* corpus:

	<i>Corpus del Español</i> (13th c.-20th c.)	<i>CORDE</i> (1950-1974)	<i>Annotated CREA</i> (1975-2000)	<i>CORPES XXI</i> (2001-2015)
<i>Fashion</i> (1490)	2 Fiction (1 instance) 1864 Oral (1 instance)	1 Miscellaneous (1 instance) 1886	-	113 Books (59 instances) 2001, 2002, 2003 2004 (3), 2005 (6) 2006 (10), 2007 (3) 2008 (3), 2009 (3) 2010 (10), 2011(14) 2012 (1), 2013 (3) Press (46 instances) 2001 (3), 2002 (4), 2003 (12), 2004 (2) 2005 (4), 2006 (3) 2007 (3), 2008 (10) 2009 (2), 2010 (2) 2011 (1) Internet (8 instances) 2006, 2007 (2) 2008, 2010 (2) 2011 (2)
<i>Fashions</i>	-	-	-	2 Books (1 instance) 2004 Press (1 instance) 2008

Table 13: Corpora distribution of *fashion* per year/register

7. Conclusions

As has been demonstrated throughout this study, the phenomenon of borrowing, understood as the linguistic consequence of a sociocultural contact between different communities, clearly shows that languages, as well as their users, are not static, but rather changeable. It is usually in their lexicons that these changes take place most quickly and are clearly seen; specifically, through language contact new foreign terms begin to be used and enter the language, sometimes making some of the original words in the recipient language fall into oblivion.

Nowadays, in our globalized world and era, the English language, unanimously recognized as the world *lingua franca*, is exerting an enormous lexical influence on many of the world languages. Furthermore, being associated with a wide array of positive connotations, English is providing contemporary Peninsular Spanish not only with terms that are really needed to name extra-linguistic realities for which Spanish as the recipient language lacks original words, but also with foreign terms which are, on the contrary, unnecessary, since the concepts and ideas they refer to can be encoded with Spanish words. As has been shown in this paper, the fashion lexicon is one particular Spanish field on which this English impact can be clearly appreciated.

The diachronic corpus-based analysis that underlies this study has revealed, in fact, that unnecessary Anglicisms in this particular field —a reflection of its advances, modernization and rapid evolution for Balteiro (2014: 159)— are becoming more recurrent in contemporary Peninsular Spanish than they have ever been. Notice, in fact, that the number of Anglicisms —44, specifically— attested in the eleven centuries that go from the tenth to the twentieth century has been much increased in the last forty-five years, from the last quarter of the twentieth century to the year 2015, totalling 661 Anglicisms.

The reasons which, I believe, are responsible for such a rise are the following ones: (i) the globalization of the world and, as a consequence, the status of *lingua franca* that English has nowadays; (ii) the highly visible presence of English, due to its selling power, in the field of advertising; (iii) and finally, the birth of new digital mass media, like the Internet, that, with their increasingly massive use as time passes, as shown in the different tables in this paper, is indeed favouring a quicker expansion of English throughout the world languages. The intermediary role that, according to Dauzat (1947: 108-111) and Pratt (1980: 224-225), among others, the different mass media —in this particular case, advertising and the Internet— play in the spread of Anglicisms to the rest of the world languages cannot, thus, be denied. It is certainly crucial to understand why the everyday contemporary Peninsular Spanish fashion lexicon is full of superfluous Anglicisms, which one day started being used in the language as part of some specific codes.²⁰

Since I firmly foresee that the presence of English terms in present-day Peninsular Spanish is not going to stop growing in the following years, the research that underlies this study should be implemented in the future in several directions to prove this hypothesis: (i) first, with a similar analysis carried out after a period of time that would cover a future time span; (ii) secondly, with a comparative study that would contrast the frequency of occurrence of the unnecessary Anglicisms that conform my corpus of study and that of their equivalent original Spanish terms; (iii) and finally, with a diachronic corpus-based analysis of some other gratuitous Anglicisms belonging to different semantic fields.

Notes

1. I would like to give special thanks to Dr. Bronwen Thomas for having hosted me as a Visiting Scholar in the *Centre for the Study of Journalism, Culture and Community* at

Bournemouth University during the summers of 2015 and 2016. There, I have had access to the Sir Michael Cobham Library, where I have found and read many of the bibliographical references cited in the present paper. I would also like to express my gratitude to Ignacio Palacios, editor of the volume, and the two anonymous reviewers who have assessed the earlier version of the manuscript, for their valuable comments and suggestions. To finish, I would like to thank my colleague, Dr. Jefferey M. Simons, for having proofread both the initial draft of the paper and its final outcome. Any errors remain my own responsibility.

2. See Gómez Capuz (2004: 19-28) for an exhaustive list of the borrowings that he calls “historic” (*prestamos históricos*), as well as of the most recent cultural borrowings, that have entered Spanish over time.

3. See, as examples, Gómez Capuz (2004: 23-25) and García Morales et al. (2016: 51-124), where two complete catalogues, organized in semantic fields, of lexical Anglicisms in contemporary Spanish are offered.

4. The dictionaries consulted are the *Diccionario Oxford español-inglés, inglés-español (DO)* (1994), the *Diccionario Larousse Español-Inglés, Inglés-Español (DL)* (1996), the *Collins Spanish-English, English-Spanish Dictionary (CD)* (1997), and the *Oxford Living Dictionary English-Spanish Online (OLD)* (2017).

5. In this paper the term “Spanish Anglicisms” (*Anglicismos hispánicos*), taken from Lorenzo (1996), refers specifically to the English borrowings attested in Peninsular Spanish. In Lorenzo (1996: 36), however, this term, described as “very ambitious”, comprises the English borrowings found in the many and diverse demographic and geographic areas where Spanish is the first official language.

6. See Gómez Capuz (1998: 150-192; 2004: 14-16) for the most significant differences between both schools and also Pulcini et al. (2012: 10-13) for a more recent and thorough discussion of the terminological diversity that exists around the phenomenon of borrowing.

7. Notice here, on the one hand, that Alfaro’s (1948) work, despite not being free of criticism, is the earliest published study of Anglicisms in Spanish, and on the other, that Pratt’s (1980) analysis is the most original and complete study of Spanish Anglicisms published up to now, thus being described by Mateescu (2006: 1) as “the decisive step towards the modern concept of Anglicisms”.

8. The classifications which Lorenzo (1996: 29) refers to are, specifically, the ones provided in Alfaro (1948), Fernández García (1972), and Pratt (1980).

9. Due to the terminological diversity previously pointed out, the distinction between direct and indirect borrowings may convey different meanings. Apart from the etymological contrasts referred to in the text which it entails, it may be used to distinguish two kinds of Anglicisms on the basis of their formal resemblance —total or partial— with the original English terms they come from (cf. Pulcini et al., 2012: 5-6).

10. The purist attitude towards necessary and gratuitous Anglicisms echoes the debate on the massive entrance of Latin terms, both required and superfluous, in the language during the Early Modern English period, known as the Inkhorn Controversy (cf. Crystal, 2004; Stehling, 2014, among others).

11. See Edwards (1994: 41) for some other reasons, apart from the mass media, that account for the worldwide spread and use of English.

12. As mentioned in Leech (2008: 174), there are many kinds of advertising: commercial advertising, trade advertising, retail advertising, prestige advertising, etc. However, since advertising refers in this study to the one “directed towards a mass audience with the aim of

promoting sales of a commercial product or service” (Leech, 2008: 174), it has to be understood as “commercial consumer advertising”.

13. I am only focusing on the main traits of verbal advertising language. For the audiovisual language in ads, see Hecker and Stewart (1988), Cook (1992), Ting-Toomey (1999), Barthes (2008), and Forceville (2008), among others.

14. In my place of residence, Huelva, a small city in the southwest of Spain, for example, even short and not specifically commercial streets are full of shopwindow signs that read “Barber’s Shop”, “Fried Chicken”, “Women’s Time” and “Carrefour Market”, among others. This same tendency is also commented on by Edwards (1994: 76), but in relation to Eastern Europe.

15. For the different pragmatic functions of Anglicisms, see González Cruz and Rodríguez Medina (2011).

16. The distinction between open and closed languages is explained by Bathia (2008: 146) in terms of their reception or resistance to borrowings. Whereas the former have been “historically receptive to the phenomenon of linguistic borrowing/foreignism”, the latter “are viewed as guardians of linguistic purity”, for they “exhibit a history of resistance to linguistic borrowing”.

17. See Davies (2010: 140) for a discussion of the size of the *CORDE* corpus.

18. The annotated version of the *CREA* corpus differs from the original non-annotated version in three main respects: (i) the number of forms it consists of: more than 126 million forms versus more than 160 million words; (ii) the linguistic varieties taken into consideration: only written language versus both written and oral language; (iii) and finally, the period of time covered: 1975-2000 versus 1975-2004.

19. Although the *Corpus del Español* provides an earlier entry date for *fashionable*, 1823, I have not considered it because it corresponds to Larra’s example — “Los **fashionables** suben y bajan a los palcos” — which in the *CREA* corpus is dated, however, as well as in Lorenzo (1996), in 1835.

20. Notice also in this regard Pountain’s (2006: 11) view, which explains syntactic borrowing as “a function of register in the sense that syntactic calques and transfers take place initially in well-defined circumscribed areas of a language”.

Primary sources (dictionaries and corpora used)

Alfaro, Ricardo J. (1948): *Diccionario de Anglicismos*. Madrid: Gredos.

CD. *Collins Spanish-English, English-Spanish Dictionary*. (1997): New York: Harper Collins Publishers.

CLAVE. *Diccionario de uso del español actual*. (1996): Madrid: S.M.

Davies, Mark (2002): *Corpus del Español*. April-June, 2017. URL: <<http://www.corpusdel-español.org>>.

DL. *Diccionario Larousse Español-Inglés, Inglés-Español*. (1996): Barcelona: Larousse.

DNLE. *Diccionario de neologismos de la lengua española*. (1998): Barcelona: Larousse.

DO. *Diccionario Oxford español-inglés, inglés-español*. (1994): Oxford: Oxford University Press.

Moliner, María (2007): *DUE. Diccionario de uso del español*. Madrid: Gredos. (Third edition)

OED. *Oxford English Dictionary Online*. (2017): September, 2017. URL: <<http://www.oed.com>>. Oxford: Oxford University Press>.

- OLD. *Oxford Living Dictionary (English-Spanish)*. (2017): May, 2017. URL: <<https://es.oxforddictionaries.com>>. Oxford: Oxford University Press.
- Real Academia Española. Banco de datos (CORDE). *Corpus diacrónico del español*. URL: <<http://www.rae.es>> [accessed April-June, 2017].
- Real Academia Española. Banco de datos (CREA, annotated version). *Corpus de referencia del español actual*. URL: <<http://www.rae.es>> [accessed April-June, 2017].
- Real Academia Española. Banco de datos (CORPES XXI). *Corpus del Español del Siglo XXI*. URL: <<http://www.rae.es>> [accessed April-June, 2017].
- Real Academia Española (2014): *Diccionario de la lengua española (DRAE)*. URL: <<http://www.rae.es>> [accessed April-June, 2017].
- Rodríguez González, Félix and Antonio Lillo Buades (1997): *Nuevo diccionario de anglicismos*. Madrid: Gredos.

References

- Balteiro, Isabel (2014): "The influence of English on Spanish fashion terminology: *-ing* Forms". *ESP Today. Journal of English for Specific Purposes at Tertiary Level*, 2(2): 156-173.
- Bartha, Éva (1997): "Aspectos psicolingüísticos del lenguaje publicitario". *Actas del XXXIII Congreso Internacional de la AEPE. Español: uniendo culturas*. Valencia: Centro Virtual Cervantes, 7-16.
- Barthes, Roland (2008): "Rhetoric of the image". In Guy Cook, ed., *The Language of Advertisement. Major Themes in English Studies*. (Vol. III). New York: Routledge, 9-22.
- Bathia, Tej K. (2008): "Discourse functions and pragmatics of mixing: Advertising across cultures". In Guy Cook, ed., *The Language of Advertisement. Major Themes in English Studies*. (Vol. II). New York: Routledge, 146-168.
- Baumgardner, Robert J. (2006): "The appeal of English in Mexican commerce". *World Englishes*, 25(2): 251-266.
- Betz, Werner (1949): *Deutsch und Lateinisch: die Lehnbildungen der althochdeutschen Benediktinerregel*. Bonn: Bouvier.
- Block De Behar, Lisa (1992): *El lenguaje de la publicidad*. Madrid: Siglo XXI. (Fourth edition).
- Bloomfield, Leonard (1933): *Language*. New York: Holt, Rinehart and Winston.
- Bonney, Bill and Helen Wilson (1990): "Advertising and the manufacture of difference." In Manuel Alvarado and John O. Thompson, eds., *The British Film Institute Media Reader*. London: British Film Institute, 181-198.
- Brennan, Timothy (1997): *At Home in the World: Cosmopolitanism Now*. Cambridge, M. A.: Harvard University Press.
- Cook, Guy (1992): *The Discourse of Advertising*. London: Routledge.
- Cook, Guy (2008a): "General introduction". In Guy Cook, ed., *The Language of Advertisement. Major Themes in English Studies*. (Vol. I). New York: Routledge, 1-19.
- Cook, Guy (2008b): "The genre of advertisement". In Guy Cook, ed., *The Language of Advertisement. Major Themes in English Studies*. (Vol. I). New York: Routledge, 237-255.
- Crompton, P. Mike and Raymund McAlea (2000): "Rhetorical devices in television advertising". In Jackie Cannon, Patricia Anne Odber de Baubeta and Robin Warner, eds., *Advertising and Identity in Europe. The I of the Beholder*. Bristol: Intellect, 32-41.
- Cronin, Michael (2003): *Translation and Globalization*. New York: Routledge.
- Crystal, David (1997): *English as a Global Language*. Cambridge: Cambridge University Press.

- Crystal, David (2004): *The Stories of English*. London: Penguin Books.
- Crystal, David (2008): "Extract from *language play*". In Guy Cook, ed., *The Language of Advertisement. Major Themes in English Studies*. (Vol. I). New York: Routledge, 417-423.
- Dauzat, Albert (1947): *La filosofía del lenguaje*. Buenos Aires: Librería el Ateneo.
- Davies, Mark (2010): "Creating useful historical corpora: A comparison of CORDE, the *Corpus del español*, and the *Corpus do portugues*". In A. Enrique Arias, ed., *Diacronía de las lenguas iberorrománicas. Nuevas aportaciones desde la lingüística de corpus*. Madrid: Iberoamericana, 139-168.
- De Mooij, Marieke (1994): *Advertising Worldwide. Concepts, Theories and Practice of International, Multinational and Global Advertising*. New York: Prentice Hall.
- Deroy, Louis (1956): *L'emprunt linguistique*. Paris: Les Belles Lettres.
- Durán Martínez, Ramiro (2002): *La lengua inglesa en la publicidad española: una aproximación pragmática*. (PhD Dissertation). Salamanca: Universidad de Salamanca.
- Dyer, Gillian (2008): "The language of advertising". In Guy Cook, ed., *The Language of Advertisement. Major Themes in English Studies*. (Vol. I). New York: Routledge, 299-315.
- Edwards, John (1994): *Multilingualism*. London: Routledge.
- Englis, Basil G. (1994): "Preface". In Basil G. Englis, ed., *Global and Multinational Advertising*. Hillsdale, New Jersey: Lawrence Erlbaum Associates, i-ix.
- Fairclough, Norman (2008): "Advertising and consumerism." In Guy Cook, ed. *The Language of Advertisement. Major Themes in English Studies*. (Vol. I). New York: Routledge, 211-222.
- Fernández García, Antonio (1972): *Anglicismos en español (1891-1936)*. Oviedo: Gráficas Lux.
- Ferraz Martínez, Antonio (1995): *El lenguaje de la publicidad*. Madrid: Arco Libros.
- Forceville, Charles (2008): "Extract from pictorial metaphor in advertisements and billboards. case studies". In Guy Cook, ed., *The Language of Advertisement. Major Themes in English Studies*. (Vol. III). New York: Routledge, 88-106.
- García Morales, Goretti, M^a Isabel González Cruz, Carmen Isabel Luján García and M^a Jesús Rodríguez Medina (2016): *La presencia del inglés en la publicidad televisiva española (2013-2015)*. Madrid: Síntesis.
- Geertz, Clifford (1986): "The uses of diversity". *Michigan Quarterly Review*, 25(1): 105-123.
- Goatly, Andrew (2008): "Advertising and consumerism". In Guy Cook, ed., *The Language of Advertisement. Major Themes in English Studies*. (Vol. II). New York: Routledge, 84-109.
- Gómez Capuz, Juan (1998): *El préstamo lingüístico: conceptos, problemas y métodos*. (Cuadernos de Filología, Anejo XXIX). Valencia: Universidad de Valencia.
- Gómez Capuz, Juan (2004): *Préstamos del español: lengua y sociedad*. Madrid: Arco Libros.
- González Cruz, M^a Isabel and M^a Jesús Rodríguez Medina (2016): "On the pragmatic function of anglicisms: A case study". *Revista Alicantina de Estudios Ingleses*, 24: 257-273.
- Gripsrud, Jostein (2002): *Understanding Media Culture*. London: Arnold.
- Halliday, Michael A. K. (1970): "Language structure and language function". In John Lyons, ed., *New Horizons in Linguistics*. Harmondsworth: Penguin, 140-165.
- Haugen, Einar (1950): "The analysis of linguistic borrowing". *Language*, 26: 210-231.
- Hecker, Sidney and David W. Stewart, eds. (1988): *Non-verbal Communication in Advertising*. Lexington, Mass: Heath and Company.
- Heldt, David, Anthony McGrew, David Goldblatt and Jonathan Perraton (1999): *Global Transformations: Politics, Economics and Culture*. Stanford: Stanford University Press.
- Hjarvard, Stig (2004): "The globalization of language: How the media contribute to the spread of English and the emergence of medialects". *Nordicom Review*, 1: 75-97.

- Jaworski, Adam (2007): "Language in the media: authenticity and othering". In S. Johnson and A. Ensslin, eds., *Language in the Media*. London: Continuum, 271-303.
- Katamba, Francis (2015): *English Words: Structure, History, Usage*. London: Routledge.
- Kelly-Holmes, Helen (2005): *Advertising as Multilingual Communication*. New York: Palgrave Macmillan.
- Klajn, Ivan (1972): *Influssi inglesi nella lingua italiana*. Florence: Olschki.
- Lázaro Carreter, Fernando (1987): "Los medios de comunicación y la lengua española". In Alonso Zamora Vicente, ed., *Primera reunión de Academias de la Lengua Española sobre el lenguaje y los medios de comunicación*. Madrid: Comisión Permanente de la Asociación de Academias de la Lengua Española, 29-43.
- Lee, Jamie Shinhee (2006): "Linguistic constructions of modernity: English mixing in Korean television commercials". *Language in Society*, 35(1): 59-91.
- Leech, Geoffrey Neil (2008): "Language and motive". In Guy Cook, ed., *The Language of Advertisement. Major Themes in English Studies*. (Vol. I). New York: Routledge, 174-179.
- Lester, Elli (1994): "International advertising research and international communication theory". In Basil G. Englis, ed., *Global and Multinational Advertising*. Hillsdale, New Jersey: Lawrence Erlbaum Associates, 3-22.
- López Morales, Humberto (1987): "Anglicismos léxicos en el habla culta de San Juan de Puerto Rico". *Lingüística Española Actual (LEA)*, 9(2): 285-303.
- Lorenzo, Emilio (1996): *Anglicismos Hispánicos*. Madrid: Gredos.
- MacGillivray, Alexander (2006): *A Brief History of Globalization: The Untold Story of Our Incredible Shrinking Planet*. London: Robinson.
- MacKenzie, Ian (2012): "Fair play to them: Proficiency in English and types of borrowing". In Cristiano Furiassi, Virginia Pulcini and Félix Rodríguez González, eds., *The Anglicization of European Lexis*. Amsterdam: John Benjamins, 27-42.
- Mateescu, Mihaela (2006): *Anglicismos en el español peninsular y americano: Criterios de clasificación*. (PhD Dissertation). Bucharest: Christian University.
- McQuarrie, Edward F. and David Glen Mick (2008): "Figures of rhetoric in advertising language". In Guy Cook, ed., *The Language of Advertisement. Major Themes in English Studies*. (Vol. I). New York: Routledge, 391-416.
- Medina López, Javier (1996): *El Anglicismo en el español actual*. Madrid: Arco Libros.
- Montes Fernández, Antonia (2006): "Globalización y publicidad: el lenguaje universal de los perfumes". *Hispanogalia*, 2: 217-226.
- Newmark, Peter (1998): *A Textbook of Translation*. New York: Prentice Hall.
- Novotná, Markéta (2007): *El Anglicismo en la lengua española*. (PhD Dissertation). Brno: Masarykova Univerzita.
- Oncins-Martinez, José Luis (2012): "Newly-coined anglicisms in contemporary Spanish. A corpus-based approach". In Cristiano Furiassi, Virginia Pulcini and Félix Rodríguez González, eds., *The Anglicization of European Lexis*. Amsterdam: John Benjamins, 217-238.
- Pennycook, Alastair (1994): *The Cultural Politics of English as an International Language*. London: Longman.
- Pennycook, Alastair (2001): "English in the world/the world in English?". In Anne Burns and Caroline Coffin, eds., *Analysing English in a Global Context*. New York: Routledge, 78-89.
- Pountain, Christopher J. (2006): "Syntactic borrowing as a function of register". In Anna Laura Lepschy and Arturo Tosi, eds., *Rethinking Languages in Contact: The Case of Italian*. Oxford: Legenda, 99-111.
- Pratt, Chris (1980): *El Anglicismo en el español peninsular contemporáneo*. Madrid: Gredos.

- Pulcini, Virginia, Cristiano Furiassi and Félix Rodríguez González (2012): "The lexical influence of English on European languages: From words to phraseology". In Cristiano Furiassi, Virginia Pulcini and Félix Rodríguez González, eds., *The Anglicization of European Lexis*. Amsterdam: John Benjamins, 1-24.
- Rantanen, Terhi (2002): *Global and the National: Media and Communications in Post-communist Russia. (Critical Media Studies)*. Lanham: Rowman and Littlefield.
- Rantanen, Terhi (2005): *The Media and Globalization*. London: SAGE.
- Robertson, Roland (1992): *Globalization Social Theory and Global Culture*. London: SAGE.
- Robles Ávila, Sara (2005): "Consideraciones sobre el léxico de la publicidad". In Sara Robles Ávila, ed., *Aspectos y perspectivas del lenguaje publicitario. (Analecta Malacitana, Anejo LVI)*. Málaga: Universidad de Málaga, 127-146.
- Rodríguez Díaz, José Antonio (2011): *Anglicismos y germanismos en el lenguaje del español peninsular contemporáneo 1998-2007*. (PhD Dissertation). Santiago de Compostela: Universidade de Santiago de Compostela.
- Rodríguez González, Félix (2016): "Prólogo". In Goretta García Morales, M^a Isabel González Cruz, Carmen Isabel Luján García and M^a Jesús Rodríguez Medina, *La presencia del inglés en la publicidad televisiva española (2013-2015)*. Madrid: Síntesis, 7-14.
- Rodríguez González, Félix (2017): "El plural de los anglicismos en español actual. Panorama y revisión crítica". *Boletín de la Real Academia Española*, 97(315): 299-329.
- Schmidt, Selina and Stefan Diemer (2015): "Comments on anglicisms in Spanish and their reception". *SWPL*, 5(8): 8-18. May, 2017. <<http://scidok.sulb.uni-saarland.de/Vol.ltexte/2015/6016/>>
- Silverstone, Roger (1999): *Why Study the Media?* London: SAGE Publications.
- Smith, Raoul N. (1982): "A functional view of the linguistics of advertising". In R. J. di Petro, ed., *Linguistics and the Professions. (Advances in Discourse Processes. Vol. 8)*. New Jersey: Ablex Publishing, 189-199.
- Stehling, David (2014): *Semantic Change in the Early Modern English Period: Latin Influences on the English Language*. Hamburg: Anchor Academic Publishing.
- Ting-Toomey, Stella (1999): *Communicating Across Cultures*. New York: The Guilford Press.
- Urrutia Cárdenas, Hernanz (2001): "La lengua en los medios de comunicación". *Onomazein*, 6: 11-30.
- Vestegaard, Torben and Kim Schroder (1985): *The Language of Advertising. (Language in Society)*. Oxford: Blackwell.
- Weinreich, Uriel (1953): *Languages in Contact: Findings and Problems*. The Hague: Mouton.
- Widdowson, Henry G. (2008): "Procedures fo discourse processing". In Guy Cook, ed., *The Language of Advertisement. Major Themes in English Studies. (Vol. I)*. New York: Routledge, 186-210.
- Williams, Kevin (2011): *International Journalism*. London: SAGE Publications.
- Yarza, Enrique (2016): "El inglés en la publicidad. Ejemplos del uso excesivo en los anuncios". *Primeras Jornadas ¿Se habla español en la publicidad?* Real Academia Española.

Analysis of the presence of Anglicisms in a Spanish internet forum: some terms from the fields of fashion, beauty and leisure

Carmen Luján García
Universidad de Las Palmas de Gran Canaria
carmen.lujan@ulpgc.es

ABSTRACT

The pervasive presence of English in Spain is unquestionable; indeed, a vast volume of literature has provided evidence of this fact. In this article, the remarkable presence of Anglicisms in a particular type of social media will be examined, namely the Spanish Internet forum *enfemenino*. The analysis covers three specific domains: beauty, fashion and leisure. The study focuses on a sample of English borrowings used in news articles published in this forum over the last 2 years (from January 2015 to March 2017). The findings reveal an increasing use of pure Anglicisms in the forum, whereas adapted Anglicisms, along with pseudo-Anglicisms, are not so common. These Anglicisms seem to be used for different reasons: the values of modernity and prestige associated with English, the lack of Spanish equivalents in some cases, the emergence of new concepts and innovations and, last but not least, the increasing influence that the Anglo-American culture is exerting on Spain. This raises the question of the extent to which these factors affect our sense of identity in Spain.

Keywords: Anglicisms, Fashion, Beauty, Leisure, Internet, Forum, Spanish

1. Introduction

The fact that English has spread as the main global language has led to question the traditional role of English merely as a foreign language (Graddol, 2006). As a consequence of this fact, we can witness the remarkable presence of Anglicisms in a



variety of languages. This expansion of English has led to the emergence and development of multiple varieties around the world, where they may be used as first, second or foreign languages. As a logical result, rather than speaking of a single English language, we can use the term “Englishes”, which was coined to refer to this set of varieties more than two decades ago by Kachru (1992: 357).

Scholars such as Görlach (2002), Anderman and Rogers (2005), Fischer and Pułaczewska (2008), De Houwer and Wilton (2011); Furiassi, Pulcini and Rodríguez-González (2012); Furiassi and Gottlieb (2015); Andersen, Furiassi and Mišić Ilić (2017) have provided insights into the contact of English with different European languages over recent decades, a period in which the influence (not to call it “invasion”) of the English language has pervaded almost every single area of daily life.

This piece of research focuses on the lexical influence of English on a Spanish Internet forum. Some previous publications have dealt with the presence of English in various fields of Spaniards’ daily life. For instance, Rodríguez González (2012), Campos (2015) and Rodríguez Medina (2016a) examined the use of Anglicisms in areas such as sports; López-Zurita (2005) showed the importance of Anglicisms in economics. Advertising and TV commercials were also investigated by Rodríguez Díaz (2011) and García-Morales, González-Cruz, Luján-García and Rodríguez-Medina (2016); the fields of eroticism and sexuality were studied by Rodríguez-González, 2011, who includes many Anglicisms in his dictionary, and Crespo-Fernández and Luján-García (2013; 2017); other studies dealt with the employment of Anglicisms in toy leaflets aimed at young children in Spain (Luján-García, 2011) or teenagers in Las Palmas de Gran Canaria (González Cruz, Rodríguez-Medina and Déniz, 2009).

The use of Anglicisms in the realms of fashion, beauty and leisure will be examined in this article. As previous studies have revealed, these fields are very likely to involve the use of Anglicisms, since the employment of foreign words provides a sense of exoticism, innovation and even creativity; in addition, all of these values are generally associated with the examined domains, namely fashion, beauty and leisure: for instance, Balteiro and Campos (2012) researched the employment of false Anglicisms in the fields of fashion and beauty, proving that in these sectors the influence of English is apparent (Balteiro and Campos, 2012: 239). After having examined 36 pseudo-Anglicisms, these authors concluded that:

the prestige and attraction of English leads speakers to follow what they consider an acceptable trend not only in terms of what they have received from English, but also through parallel coinages; the only justification for which probably is their “Englishness” – that is, the iconic power of English as a trendsetter (Balteiro and Campos, 2012: 249).

Other pieces of research that have dealt with fashion-related topics may also be referred to: Rodríguez-Medina (2016b) also studied the use of Anglicisms in TV commercials of cosmetics, hygiene and personal care products. This author analysed 531 TV commercials across four Spanish TV channels (Tele5, Antena3, LaSexta and Disney Channel) and she found that:

[t]he results confirm a considerable presence of pure Anglicisms, English-Spanish code switching, pseudo-Anglicisms and Anglo-American imagery and music in the advertising of products related to cosmetics, hygiene and personal care on Spanish television (Rodríguez-Medina, 2016b: 157).

González-Cruz (2015) explored the use of Anglicisms in several leisure fields: technology, entertainment, and food and drinks in the TV commercials of three Spanish TV channels: she found that the English language and the Anglo-American culture have a great impact on the Spanish language and culture with regard to leisure-related topics. Díez-Arroyo (2016) focused her research on ten special issues of fashion magazines, including *Elle*, *Hola*, *Telva* and *Vogue* among others, published between March and September 2013 and she asserted that:

Spanish fashion magazines regard stylistic choices as a persuasive strategy to reach and appeal to their wide readership. Journalists have found in Anglicisms the perfect elements to perform this rhetorical function (Díez-Arroyo, 2016: 38).

The means of communication that will be examined in this article is the Internet forum www.enfemenino.com. It is beyond doubt that Europeans use the Internet on a daily basis, and their exposure to this means of communication is not only work-related, but it is also used as a source of entertainment. Discussion forums are a perfect way to interact with other people, as well as to share or exchange information and publish news, among other activities. Despite the numerous studies that have been referred to above, so far not many pieces of research have focused on the analysis of Anglicisms in the particular context of Internet forums: for instance, Garley and Hockenmaier (2012) reported on the use of Anglicisms in a German hip hop forum; Zhang (2015) examined multilingual creativity in a Chinese microblog, *Shanghai Release*, involving English among other languages; Crespo-Fernández (2015) examined the use of taboo and euphemistic words in some Internet forums; Tagliamonte (2016) has recently studied the linguistic uses of the Internet by North American youth focusing on different linguistic aspects, such as acronyms and intensifiers. All these studies have contributed to shedding some light on this fascinating field of research, which is constantly developing as social media technologies expand in the contemporary world. New trends and ways to interact online emerge daily, and in Tagliamonte's and Denis's (2008: 27) words, "[a]ll these provide yet-to-be-discovered venues in which the foremost commodity is language".

This study intends to bridge a gap in the literature by analysing Anglicisms in a Spanish Internet forum, especially focusing on their use in the domains of fashion, beauty and leisure. The initial hypotheses are the following:

- the specialized language of fashion, beauty and leisure tends to include Anglicisms as a resource to attract the audience and provide a sense of modernity and coolness to forum users;

- pure Anglicisms, namely English borrowings that have not undergone any kind of adaptation to the recipient language – Spanish, in this case, are the most frequently employed in the examined fields;
- fashion is probably the area with the largest use of all the Anglicisms examined.

The main objectives of this analysis intend to provide a reply for the following research questions:

- What type of Anglicism (pure, adapted, false or hybrid) is the most frequently employed in the examined Spanish Internet forum?
- In which subject area (fashion, beauty or leisure) is the use of Anglicisms most prolific?

2. Methodology

The methodology employed in this analysis is based on a careful reading of each news article published in an online forum. After that, a manual compilation of those posts which contained any kind of Anglicism was carried out. Therefore, the collection of the sample was quite laborious, but this process was considered to be the most appropriate and accurate. Once the sample was compiled, Anglicisms were selected and examined using the following method: firstly, the different types of Anglicisms encountered were classified according to the categorisation that will be presented in the following lines of this section, and, secondly, the domain in which the Anglicisms were used was identified with reference to the context of the post as a whole. The three domains that were distinguished were fashion, beauty and leisure.

Some Anglicisms were used more than once in the same news article, therefore, not only the variety but also the frequency of use of the chosen Anglicisms is examined and displayed in the tables present in the appendix (see Tables 4, 5, 6 and 7). Quantitative and qualitative analyses are presented in section 3.

The forum chosen for this analysis was *enfemenino*, available at www.foroenfemenino.com, whose main audience and participants are made up of women, as its name suggests. However, contributions are not restricted to the female gender – men can also take an active part in it. This forum belongs to the international group *aufemininS.A.* a media company majority-owned by German media group Axel Springer. According to data offered by this website in March 2017, this group had more than 14 million readers and had 420 million pageviews in the U.S. alone in January 2017.

The contributions to this forum are in Spanish, and its team publishes articles about fashion, beauty, leisure, society, cooking, weddings, home, travelling, cinema and TV, and games. In this study, the focus is on some of these sections, namely fashion, beauty and leisure; In addition, rather than analysing the contributors' posts, this piece of research focuses on the regular publications by the journalists or contributors from the Spanish-speaking team of this website. The period of compilation of the corpus ranged

from 1st January 2015 to 31st March 2017, which encompasses a period of two years and three months.

As above mentioned, this piece of research focuses on the analysis of Anglicisms in some particular domains, so for the purpose of this research, an Anglicism is defined as any term that is directly borrowed from English. Different categorisations of Anglicisms have been suggested (Alfaro, 1970; Lope Blanch, 1977; Pratt, 1980; Lorenzo, 1987), but the most recent one is by Furiassi, Pulcini and Rodríguez-González (2012). Their widely accepted typology distinguishes adapted, non-adapted/pure Anglicisms, false/pseudo-Anglicisms, hybrid Anglicisms and calques. However, in the context of this study, this categorisation must be adapted to suit the corpus compiled. Consequently, in this study the following types of Anglicisms are considered:

Non-adapted or pure Anglicisms: direct Anglicisms including “a word or a multi-word unit borrowed from the English language with or without minor formal or semantic integration, so that it remains recognizably English in the recipient language (RL)” (Pulcini et al., 2012: 6).

Adapted Anglicisms: words or multi-word units borrowed from English with orthographic, phonological and/or morphological integration into the structures of the RL. Both terms, the source language (SL) term and the recipient language (RL) term, are close in meaning (Pulcini et al., 2012: 7). Therefore, adaptations tend to affect the morphology rather than the meaning of the words.

False or pseudo-Anglicisms: the definition of false or pseudo-Anglicism considered for this analysis is the following:

a word or idiom that is recognizably English in its form (spelling, pronunciation, morphology, or at least one of the three), but is accepted as an item in the vocabulary of the receptor language even though it does not exist or is used with a conspicuously different meaning in English (Furiassi, 2010: 34).

Hybrid Anglicisms: a mixture or combination of two words from different languages, one being necessarily English.

In the following lines, the compiled list of Anglicisms is broken down according to the three different subject areas investigated: fashion, beauty and leisure. It is worthwhile to clarify that, in some cases, it is difficult to set clear-cut boundaries between beauty and fashion. The following criteria are applied in order to distinguish these three different semantic fields:

- 1) *fashion Anglicisms* include Anglicisms related to clothes, trends, kinds of garments and shoes, accessories and home decoration;
- 2) *beauty Anglicisms* encompass Anglicisms related to beauty treatments, hairstyle, make-up and nail decoration;
- 3) *leisure Anglicisms* include Anglicisms related to free-time activities and hobbies.

3. Findings

This analysis initially focuses on the breakdown of the different types of Anglicisms listed in the previous section. Firstly, a quantitative analysis will provide general figures and, after that, a qualitative analysis will examine the findings in detail. In the appendix, Tables 4, 5, 6 and 7 display all the Anglicisms under scrutiny in this analysis.

Table 1 below summarises the frequency and percentage of the different types of Anglicisms detected in *enfemenino* during the examined period.

TYPES OF ANGLICISMS	FREQUENCY	%
Non-adapted or pure	66	83.5
Adapted	3	3.8
Pseudo- or false	8	10.1
Hybrid	2	2.6
Total	79	100

Table 1: Types and frequency of Anglicisms in *enfemenino*

When it comes to the frequency of the Anglicisms compiled in the sample, it was found that most of them are only used once, twice or three times. However, there are a few that are used with a higher frequency. This is the case for the following Anglicisms, listed in Table 2.

ANGLICISMS	FREQUENCY	%
<i>look</i>	47	58
<i>afterwork</i>	11	13.6
<i>celebrities/celebs</i>	11	13.6
<i>dot eyeliner</i>	7	8.6
<i>phubbing</i>	5	6.2
Total	81	100

Table 2: The five most frequent Anglicisms in *enfemenino*

The reason why *look* is by far the most frequently used Anglicism could be due to the fact that there is a special weekly section in the analysed forum which deals with the worst look of the week. In second position, there are *afterwork* and *celebrities/celebs*, both of them representing 13.6% of the total number of the five most frequent Anglicisms. *Dot eyeliner* is in fourth position, with 8.6%, and in the fifth and last position ranks *phubbing*, with 6.2%.

Table 3 breaks down the different Anglicisms used within the three examined domains. This table also provides the frequency of use of Anglicisms within these three domains.

DOMAIN	VARIETY	%	FREQ.	%
fashion	34	43.6	100	49
beauty	23	29.5	49	24
leisure	21	26.9	55	27
Total	78	100	204	100

Table 3: Different Anglicisms and frequency of Anglicisms by domain in *enfemenino*

Figure 1 below shows the different Anglicisms in the domains of fashion, beauty and leisure.

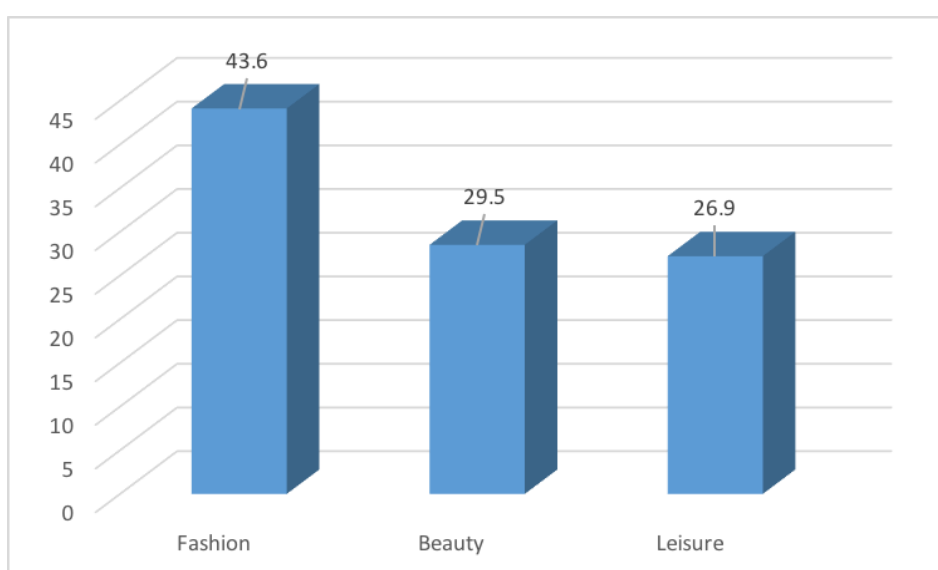


Figure 1: Different types of Anglicisms in fashion, beauty and leisure

As shown in Table 3 and Figure 1, the domain of fashion is the one with the largest variety of Anglicisms with 43.6% of the total, beauty being in second position with 29.5% and leisure in the third and last position with 26.9%. Figure 2 shows the frequency of use of Anglicisms in the examined domains.

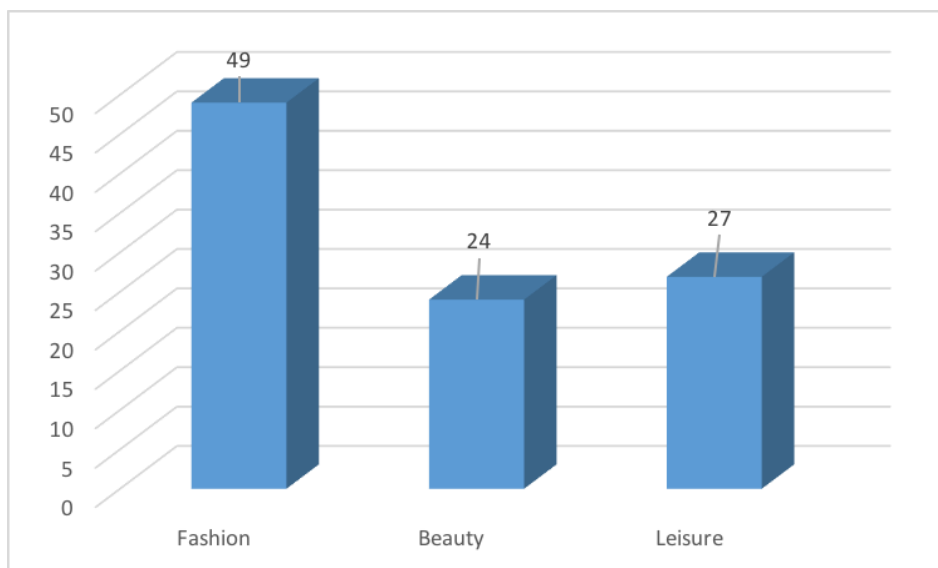


Figure 2: Frequency of Anglicisms in fashion, beauty and leisure

As Table 3 and Figure 2 display, when it comes to the frequency of Anglicisms by domain, fashion is again in first position, with a total of 100 occurrences, i.e. 49%, followed by leisure items, with 55 occurrences, i.e. 27%, and eventually beauty-related terms, with 49 occurrences, i.e. 24%.

It should be mentioned that Anglicisms are used in a variety of ways in the Spanish forum examined. In some cases, Anglicisms are placed between inverted commas; at times italics is used; in other cases, no particular punctuation or font is used. For example, the term *baking* occurs in italics, as shown in the following example.

(1) Por ello, cuando oímos que había un nuevo método de maquillaje llamado *baking*, nos temimos lo peor y nos imaginamos a mujeres "horneando" su cara para verse más guapas (21/07/2015).

On the contrary, the term *candy* is placed between inverted commas in order to highlight that it is a foreign term, as shown in the following example.

(2) El **estilo nórdico** es el que más se presta a adoptar estos tonos y es que la combinación de los básicos blanco y negro con los "candy" da como resultado unos ambientes perfectos y muy acogedores (05/11/2015).

As anticipated, there may be Anglicisms that are not marked at all. These are the cases of *outfit* and *look* in the following example:

(3) Entre ellas encontrarás a Paris Hilton con un look de exploradora, a **Elle Fanning** con su outfit para el gimnasio o a **Zoe Kravitz** con una elección que no le favorece. ¡Echa un vistazo! (17/03/2017).

This lack of consistency in terms of devices used to highlight the foreign nature of these foreign lexical units reveals that many of the examined Anglicisms are not totally assimilated¹ in Spanish.

In the following lines, a qualitative analysis will be carried out to examine in depth some of the Anglicisms found in the sample. Despite the fact that in the appendix an example of the use of these Anglicisms in context is provided, some more examples will be presented in the subsequent sections. The Anglicisms included in these examples are marked in bold type with the aim of highlighting their use, but they do not necessarily occur in bold in the original post, as commented in the previous lines.

3.1. Pure Anglicisms (66 cases)

The following list of pure or non-adapted Anglicisms compiles a total of 66 examples, some of which present different spellings.

After party, baking, beauty, beauty youtubers, candy, celebrity/celebrities/celebs, checkout, chokers, clown contouring, cookies, cottage, contouring, dot eyeliner, dress code, eyeliner, foodie, front row, gif, girls' night out, gloss, glowing hair, greenery, grills, holographic lips, influencer, in-flight, jeans, layering, look, makeup, mix, nail art, nude, outfit/outfits, over size, performance, phubbing, piercing, pink lady, pink lips, pop-up stores, rainbow food, rainbow hair, rainbow freckless, reality, room service, sane food, selfie, sexy, shopping, shorts, skyline, slim, slip dress, soft, spa, streetstyle, sweet, teenager, tip, top, trendy, tweed, welcome pack, wire nail, and working girl.

The case of *after party* can be regarded as a spaced compound, since it is written with a space between the preposition *after* and the noun *party*. The Anglicism *beauty* is also noticeable, since there is a direct equivalent in Spanish for this term, namely *belleza*, but in four cases the English word is chosen, as shown in the following post.

(4) Y si hoy no sabes cómo maquillarte, pregúntale a nuestro test y él te resolverá todas tus dudas **beautyen** un solo click (03/01/2017).

Candy is another lexical unit that deserves some attention, as it refers to a decorating style involving a mixture of different colours, frequently related to hair or even walls, as in the following example.

(5) El estilo nórdico es el que más se presta a adoptar estos tonos y es que la combinación de los básicos blanco y negro con los “**candy**” da como resultado unos ambientes perfectos y muy acogedores. (05/11/2015).

Celebs is another remarkable case as it is the widely accepted abbreviation of the English term *celebrities*: although it is a variant of the same word, it has been counted as an Anglicism per se. The term *celebrity* and its plural form *celebrities*, along with the abbreviated form *celebs*, is one of the most frequently used, and it refers to famous people

who are able to influence and even change fashion and beauty trends. Obviously, in Spanish there is the equivalent *famosos*, but in the area of fashion, the English term is sometimes preferred, since it provides a sense of modernity and coolness that the Spanish term does not have.

(6) Las Semanas de la Moda marcan las tendencias de la nueva temporada y dejan a su paso un montón de looks de **celebrities** que pasarán a la historia (24/02/2017).

Contouring is a style of makeup, first used by Kim Kardashian, which slims down the face by using different tones on the skin. The following examples, 7 and 8, provide evidence of the context in which the Anglicism is used.

(7) A new style, called **clown contouring**, has emerged. Después de la fiebre del contorneado que inició Kim Kardashian hace un par de años, y del nuevo método del **clown contouring** (21/07/2015).

(8) Ahora esta tendencia ha dado un paso más, el llamado **clown contouring**, nada más y nada menos que incluir colores llamativos a tu maquillaje, y divertirse maquillándose como si fueras un payaso, de ahí su nombre. El proceso de maquillar puede resultar muy tedioso, así que ¿por qué no divertirse mientras le das forma a tu cara? Aunque a primera vista pueda parecer una locura, el resultado de este **clown contouring** es asombroso, ya que tras difuminar los dibujos, el maquillaje tiene multitud de matices (16/07/2016).

Dot eyeliner is a new trend in makeup, which consists of drawing a small dot in the line of the eyelashes. Example 9 shows a post in which this term is used.

(9) Emma Roberts nos tiene acostumbradas a ser una fiel seguidora de las últimas tendencias tanto en el mundo beauty como en el mundo de la moda así que, no nos ha sorprendido en absoluto que haya comenzado a usar el **dot eyeliner** (31/01/2017).

Foodie is an informal epithet used to talk about a person who is fond of food. It has become popular in Spanish, probably after the growing interest in TV shows dealing with food, such as *Masterchef*, just to mention one example.

(10) Este artículo es para ti, querida amiga **foodie** (27/01/2017).

Look is one of those Anglicisms which have been accepted and included in the *Diccionario de la Lengua Española (DLE)* and *Diccionario Panhispánico de Dudas (DPD)*. It is quite widespread, as has been used in Spanish for a long time. According to Rodríguez González (2017: 612), this term started being assimilated in Spanish towards the middle of the last century: despite the existence of the equivalents *imagen* and *apariencia*, this Anglicism is quite frequent in Spanish. That is the reason why this lexical unit shows the highest frequency of use.

The following examples 11 and 12 provide evidence of the use of *look* in some posts.

(11) Tenemos que hablar seriamente del **look** de Lena Dunham en la fiesta de despedida de la serie *Girls* (10/02/2017). [...] Esta vez y a pesar de la gran pérdida de peso por la que está pasando, no le hemos encontrado la gracia a su **look** (10/02/2017).

(12) Kendall Jenner se viste como Caperucita blanca para salir de fiesta por Londres y añade un diente de oro para completar un **look** horrible (24/03/2017).

Phubbing is a neologism in English, a blend of *phone* and *snubbing*, meaning ‘the practice of ignoring one’s companion or companions in order to pay attention to one’s phone or other mobile device’ (*OLD*). This term was first attested in 2012 coined by an Australian advertising agency as part of a marketing campaign with the Macquarie Dictionary; the next example (13), shows how this lexical unit has been very rapidly introduced into Spanish. In our list of examples, *phubbing* is attested in 2016, so in a period of four years, the term is being used in Spanish.

(13) **Phubbing**, la tendencia que practicas sin darte cuenta y que pone en peligro tus relaciones. Podríamos decir que en la última semana has sufrido al menos dos o tres casos de **phubbing** sin saber que estaban teniendo lugar (22/08/2016).

Outfit is an Anglicism that may replace the Spanish terms *ropa* or *vestimenta*. The English term is endowed with fashionable nuances, as shown in examples 14 and 15:

(14) Natalia Vodianova ha cometido un crimen estilístico esta semana. Descubre quién es el culpable de que la modelo vistiera de esta guisa y por qué nos ha defraudado tanto su **outfit** (27/01/2017).

(15) Entre ellas encontrarás a Paris Hilton con un look de exploradora, a Elle Fanning con su **outfit** para el gimnasio o a Zoe Kravitz con una elección que no le favorece. ¡Echa un vistazo! (17/03/2017).

Outfits, in the plural form, is also used in English as a direct translation of *las prendas*, as shown in example 16.

(16) ¡llega el momento de preparar los **outfits**! (03/01/2017).

After having examined in detail some examples of pure Anglicisms, it could be asserted that this is the type where most cases have been found. A recent piece of research by García-Morales et al. (2016) demonstrates that pure Anglicisms, without any kind of adaptation, are used more frequently than other types (adapted, false, hybrids, calques, among others) in Spanish.

3.2. Adapted Anglicisms (2 cases)

The terms in this category have all undergone some kind of adaptation in Spanish. Indeed, *clímax* (*E.* < *climax*), and *tartán* (*E.* < *tartan*), both show orthographic adaptation.

Firstly, in the case of *clímax* a stress mark on the vowel *-i* is added to adapt the term to Spanish graphemic conventions. In the case of *tartán* (*E.* < *tartan*), which is a style of cloth of Scottish origin with a pattern of different coloured straight lines crossing each other, an adaptation may be observed, since a stress mark is added to the last *-á-*, as Spanish words ending in *-n* require. See example 17.

(17) **Tartán** pero en total look (20/08/2015).

3.3. Pseudo-Anglicisms (8 cases)

The use of Pseudo-Anglicisms seems to be common in the field of fashion journalism and motivated by stylistic purposes, i.e. to have an impact on the audience (Furiassi, 2010: 62).

Afterwork, *bombers* and its stressed variety *bómber*, *fashion*, *reality*, *teenager*, *short* and *sport* are included in this category. These examples may be regarded as false Anglicisms since they do not exist in English, at least with the meaning in which they are used in Spanish.

The term *afterwork*, used in eleven cases in the examined corpus, is frequently employed in Spanish to refer to that drink that you have with your work mates after work and with the intention of getting to know other workmates a step further from the job context. Whereas in English *afterwork* is used as an adjective, in expressions such as “after work drinks”, in Spanish it is used as a noun: this borrowing could be regarded as a metonymic semantic change in Rodríguez-González’s (2013: 135-136) classification of pseudo-Anglicisms, since its meaning has been extended to the *afterwork club* or *bar* where people meet their colleagues. Example 18 shows the use of this term in context:

(18) Organizar un **afterwork** con tus compañeros de trabajo puede resultar súper positivo a la hora de estrechar la relación con ellos. Los **afterwork** tienen la maravillosa capacidad de hacernos mucho más llevadera la semana, lo prometemos (02/12/2015).

The reason why the borrowing *bómber* is considered a pseudo-Anglicism is because the actual English form is *bomber*, a noun used with an adjectival function that usually modifies the noun *jacket*, as in *bomber jacket*. In Spanish, *bombers* is the plural of *bomber*, meaning *bomber jackets*; in English *bombers* refers to any animate or inanimate object that can carry bombs, e.g. bomber airplanes. This style is inspired in the design of the jackets that pilots used to wear during World War II. As shown in posts (19) and (20), both spellings are used in the examined posts: *bomber* and *bómber*.

(19) Lleva tu **bómber** como Gigi Hadid y Kendall Jenner (08/02/2017).

(20) Calcetines de rugby, **bombers**, sweaters y abrigos acolchados te vestirán de arriba abajo (29/07/2016).

Fashion has been included in this category of false Anglicisms, as is used as an adjective rather than a noun, as in the example reported (21), being the real English equivalent 'fashionable'.

(21) La Kardashian Jenner más **fashion**, sencilla y natural ha vuelto a meter la pata (24/02/2015).

Reality is used in the examined posts as the elliptical form of the English compound term *reality show*, as the following example (22) displays.

(22) Tras el éxito de su **reality**, mamá Jenner no quiso dejar pasar la ocasión y materializó la belleza de sus hijas mayores (14/06/2016).

Another case is represented by *teenager*, which is employed as the English adjective *teenage*, as in the following illustration (example 23).

(23) Todo muy **teenager**, para tener contento a su público. (14/06/2016)

Short, without the final *-s*, is used to talk about shorts or hot pants. In Spanish, this use is quite common, whereas in English *short* is just an adjective. Example 24 provides evidence of this use.

(24) No podría faltar el ya mencionado **short** en satén fucsia, la camisa estampada, la cazadora de ante con mangas de pelo... (28/10/2016).

The false Anglicism *sport* is employed in Spanish to refer to a sporty style. Indeed, in Spanish this noun is often used with the function of the English adjective *sporty*, as shown in example 25.

(25) Eso sí, hay que tener en cuenta el diseño porque si es muy **sport** serán más adecuadas para looks informales (09/02/2017).

Beautiful box, which is not included in the analysis as it is a proper noun, despite not being part of the study, is a marked case which is used as a proper noun, and it refers to a parcel that any subscriber will receive monthly for less than 15.90 Euros, would indeed be considered a false Anglicism. It contains beauty products and it is aimed at women of all ages. In Spanish, the expression *beauty case* (not present in this study) is also used as a pseudo-Anglicism. Instead, in English the term which is used is *vanity case* (Furiassi, 2010; Balteiro and Campos, 2012).

(26) Si deseas suscribirte a **Beautiful Box** pero estás esperando la mejor oferta, estás de enhorabuena. Descuento de un 5% en tu **Beautiful Box** (01/01/2016).

As Rodríguez-González (2013: 147) reports, many pseudo-Anglicisms are relatively well-established in Spanish, and that is the reason why they have been included in

different dictionaries. This is the case for *short* and *sport*, both of them present in this study and recorded in the *Diccionario de la Lengua Española (DLE)* and the *Diccionario Panhispánico de Dudas (DPD)*.

3.4. Hybrid Anglicisms (2 cases)

Country chic is an English-French hybrid, which is used to refer to a particular style. Example 27 illustrates this combination of the English term *country* and the French word *chic*.

(27) Piezas chics que construyen los detalles de este estilo evocando los ambientes más románticos y bucólicos acercándose incluso al **country chic** (18/08/2015).

Ombre lips is another instance of hybrid Anglicism since it combines the French term *ombré* with the English term *lips*. Examples 28 and 29 provide evidence of this hybrid in context.

(28) ¿Eres de las que no sale de casa sin llevar los labios pintados pero te apetece probar algo diferente? ¿Hay días en los que no te decides entre un labial u otro? ¡Ponte los dos a la vez! La tendencia **ombre lips** vuelve con fuerza este verano (19/06/2016).

(29) Si esta es tu primera vez con los **ombre lips**, lo mejor es que pruebes con un degradado suave, dentro de la misma tonalidad de color. Una vez que ya tengas más práctica podrás probar con otros más atrevidos, que mezclen colores totalmente distintos (19/06/2016).

The following sub-sections will break down the corpus considering the three different sectors described, namely fashion, beauty and leisure.

3.5. Fashion Anglicisms (33 cases)

As reported by Lopriore and Furiassi (2015: 200), “the field of textiles and materials is highly technical, referential, exact, denotative and monosemous, with a limited communicative cycle. It is marked by borrowings from other languages.”

The pure Anglicisms that fit within this category are the following: *candy*, *chokers*, *cottage*, *dress code*, *front row*, *grills*, *jeans*, *layering*, *look*, *mix*, *nude*, *outfit*, *over size*, *sexy*, *shorts*, *slim*, *slip dress*, *soft*, *streetstyle*, *sweet*, *tip*, *top*, *trendy*, *tweed*, *working girl*. There are also some adapted Anglicisms: *bómbier* and *tartán*. Some false or pseudo-Anglicisms also appear within this category: *bombers*, *fashion*, *short*, *sport* and *teenager*. *Country chic* is a hybrid Anglicism within this realm.

3.6. Beauty Anglicisms (23 cases)

In this subsection, the following pure Anglicisms are included: *baking, beauty, beauty youtubers, clown contouring, contouring, dot eyeliner, eyeliner, gloss, glowing hair, greenery, holographic lips, in-flight, influencer, makeup, nail art, piercing, pink lady, pink lips, rainbow hair, rainbow freckless* and *wire nail*. *Ombre lips*, as explained above, is a hybrid Anglicism, as it combines a French term with an English term.

3.7. Leisure Anglicisms (21 cases)

The following list encompasses those Anglicisms that have been considered as belonging to the field of leisure since they are used to refer to free time and entertainment activities. These Anglicisms are *afterwork* (used as a noun, despite being used as an adjective like in English), *after party, celebrity/celebrities/celebs, check out, cookies, foodie, gif, girls' night out, performance, phubbing, pop-up stores, rainbow food, room service, sane food, selfie, shopping, skyline, spa, welcome pack*. There is also one case of an adapted Anglicism in this area, and one case of false Anglicism, *reality*.

In these specialised areas, the use of Anglicisms has different functions, as Rodríguez González (1996)² reported two decades ago. The referential function, which denotes the literal meaning of a term, when applied to the use of Anglicisms in Spanish, may intend to fill a lexical gap whenever there is no suitable equivalent available. This is the case for lexical units like *selfie*.

The expressive function includes stylistically marked lexical units that normally have an emotive connotation. These Anglicisms usually have positive or negative connotations. Some words taken from this corpus could be regarded as fulfilling the function of attracting the audience, looking fashionable, cool and even snobbish. For example, *look* instead of *apariencia*; *outfit/outfits* instead of *prendas*, *shorts* rather than *pantalones cortos*, *teenager* instead of *adolescente*. All these Anglicisms, to list but a few, imply positive or fashionable nuances when used in Spanish.

Finally, Rodríguez González (1996) distinguishes another function; the textual one, or the capacity the language has to create text in relation to the context. For example, for the economy of language, as in the case of *jeans* to avoid saying *pantalones vaqueros*.³

4. Conclusion

This study is intended to reveal current trends about the use of Anglicisms in a specialized Spanish forum aimed mainly at women. After having carried out this analysis, the first hypothesis is confirmed: there is a tendency to use Anglicisms in the domains of fashion, beauty and leisure. The second hypothesis is also confirmed: non-adapted or pure Anglicisms, 66 in total, are by far more widespread than other types: false Anglicisms rank second, and adapted Anglicisms and hybrid Anglicisms last.

In terms of the subject area most likely to be affected by the use of Anglicisms, it is important to state that, due to the type of sources used for this research, in some cases, it

is difficult to draw a line between beauty and fashion. The whole context of the post provided the clue to discern whether to include some Anglicisms in the fashion or beauty domains. However, as the third initial hypothesis predicted, fashion seems to be the area where most Anglicisms are found, followed by beauty and then leisure.

In a globalised world, hot-off-the-press English neologisms, such as *phubbing*, for instance, are immediately exported to other languages, without leaving sufficient time for proper translation. In addition, “the taste for the exotic” and “the charm of a foreign language” (Furiassi, 2010: 63) are important motivations for the use of Anglicisms, especially in domains such as fashion, beauty and leisure, which are intrinsically subject to constant change. French was traditionally the language associated with fashion until the 1980s, when English started replacing French as the language most frequently used to talk about fashion-related topics (fashion magazines, TV channels, social networks) (Lopriore and Furiassi, 2015: 203).

Despite the growing number of publications (Balteiro, 2011; Lopriore and Furiassi, 2015) dealing with Anglicisms in the specialised language of fashion, little attention has been paid so far to the analysis of the domains of beauty, fashion and leisure in online forums. The main purpose of this article has been to bridge this research gap and shed some light on the actual use of Anglicisms in this online medium, although its limitation lies in the analysis of one single Internet forum.

It is unquestionable that the contact between English and Spanish represents a cultural and linguistic enrichment for the recipient language, Spanish in this case. These linguistic changes, mostly visible at the lexical level, may also bring cultural changes in the way people understand and express certain concepts by using English rather than Spanish. The adoption of these lexical innovations mirrors the acceptance of cultural patterns typical of the donor language and culture. Therefore, the findings of this study suggest some reflections on the degree of exposure to the English language and the Anglo-American culture on the part of Spanish speakers. More precisely, these domain-specific Anglicisms may exert a noticeable influence and have subsequent long-term effects on the Spanish sense of identity. As Bloomfield (1933: 445-458) asserted, “every speech community learns from its neighbors” and, consequently, “cultural loans show what one nation has taught another”.

As far as future research is concerned, it would be of utmost interest to analyse other Internet forums in order to check whether this tendency is similar in different Spanish forums.

Notes

1. When a term is assimilated, in Lorenzo’s (1987) terminology, refers to a word has been completely accepted and naturalised in the recipient language, Spanish in this case, long time ago. Many of these terms have been included by DLE or are widely accepted and used by speakers of that RL.

2. Rodríguez González (1996) proved that Anglicisms in Spanish generally respond to specific functions, namely referential, expressive and textual.

3. As Rodríguez González (1996) stated in his categorisation, it is essential to notice that many of the lexical units compiled in this sample may fulfil more than one function.

4. I would like to thank the two anonymous reviewers who have given me valuable feedback on a previous version of this manuscript. Their useful and insightful comments have helped to improve this article.

References

- Alfaro, Ricardo (1970): *Diccionario de anglicismos*. Madrid: Gredos.
- Anderman, Gunilla, and Margaret Rogers, eds. (2005): *In and Out of English: For Better, For Worse?* Clevedon: Multilingual Matters.
- Andersen, Gisle, Cristiano Furiassi and Biljana Mišić Ilić, eds. (2017): *The Pragmatics of Borrowing: Investigating the Role of Discourse and Social Context in Language Contact* [Special Issue]. *Journal of Pragmatics*, 113. Amsterdam: Elsevier.
- Balteiro, Isabel (2011): "A few notes on the vocabulary of textiles and fashion". In I. Balteiro, ed., *New Approaches to Specialized English Lexicology and Lexicography*. Newcastle upon Tyne: Cambridge Scholars Publishing, 65-81.
- Balteiro, Isabel and Miguel Ángel Campos (2012): "False anglicisms in the Spanish language of fashion and beauty". *Ibérica*, 24: 233-260.
- Bloomfield, Leonard (1933): *Language*. New York: Henry Holt.
- Campos-Pardillos, Miguel Ángel (2015): "All is not English that glitters: false anglicisms in the Spanish language of sports". *Atlantis*, 37(2): 155-174.
- Crespo-Fernández, Eliecer (2015): *Sex in Language: Euphemistic and Dysphemistic Metaphors in Internet Forums*. London: Bloomsbury.
- Crespo-Fernández, Eliecer and Carmen Luján-García (2013): "Anglicismo y tabú: valores axiológicos del anglicismo". *Revista de Estudios Filológicos*, 52: 53-74.
- Crespo-Fernández, Eliecer and Carmen Luján-García (2017): "Anglicisms and word axiology in homosexual language". *Revista Española de Lingüística Aplicada (RESLA)*, 30(1): 74-102.
- De Houwer, Annick and Antje Wilton, eds., (2011): *English in Europe Today*. Amsterdam and Philadelphia: John Benjamins.
- Díez-Arroyo, Marisa (2016): "English words as euphemisms in Spanish fashion". *English Today*, 127(32/3): 30-39.
- Fischer, Roswitha, and Hanna Pułaczewska, eds. (2008): *Anglicisms in Europe. Linguistic Diversity in a Global Context*. Newcastle Upon Tyne: Cambridge Scholars Publishing.
- Furiassi, Cristiano (2010): *False Anglicisms in Italian*. Monza: Polimetrica.
- Furiassi, Cristiano, Virginia Pulcini and Félix Rodríguez-González, eds. (2012): *The Anglicization of European Lexis*. Amsterdam and Philadelphia: John Benjamins.
- Furiassi, Cristiano and Henrik Gottlieb, eds. (2015): *Pseudo-English: Studies on False Anglicisms in Europe* [Language Contact and Bilingualism 9]. Boston & Berlin: De Gruyter Mouton.
- García-Morales, Goretti, Isabel González Cruz, Carmen Luján García and M^a Jesús Rodríguez Medina (2016): *La presencia del inglés en la publicidad televisiva española (2013-2015)*. Madrid: Síntesis.
- Garley, Matt and Julia Hockenmaier (2012): "Beefmoves: Dissemination, diversity, and dynamics of English borrowings in a German hip hop forum". *Proceedings of the 50th Annual Meeting of the Association of Computational Linguistics*. Jeju: Republic of Korea: 135-139. Available at <http://www.aclweb.org/anthology/P12-2027> [accessed 29 September 2017].

- González Cruz, M^a Isabel, M^a Jesús Rodríguez-Medina and M^a Jesús Déniz-Santana (2009): *Los anglicismos en el habla juvenil de Las Palmas de Gran Canaria*. Madrid: La Factoría de Ediciones.
- González Cruz, M^a Isabel (2015): “Anglicising leisure: The multimodal presence of English in Spanish TV adverts”. *Calidoscopio*, 13(3): 339-352.
- Görlach, Manfred, ed. (2002): *English in Europe*. Oxford: Oxford University Press.
- Graddol, David (2006): *English Next*. Plymouth: British Council.
- Kachru, Braj (1992): “Teaching world Englishes”. In B. Kachru, ed., *The Other Tongue: English across Cultures*. Chicago: University of Illinois Press, 355-366.
- Lope Blanch, J.M (1977): “Anglicismos en la norma lingüística culta de México”. In J.M. Lope Blanch, ed., *Estudios sobre el español hablado en las principales ciudades de América*. México: UNAM, 271-279.
- López-Zurita, Paloma (2005): “Economic anglicisms: Adptation to the Spanish linguistic system”. *Ibérica*, 10: 91-114.
- Lopriore, Lucilla and Cristiano Furiassi (2015): “The influence of English and French on the Italian language of fashion: Focus on false Anglicisms and false Gallicisms”. In C. Furiassi and H. Gottlieb, eds., *Pseudo-English: Studies on False Anglicisms in Europe* [Language Contact and Bilingualism 9]. Boston & Berlin: De Gruyter Mouton, 197-226.
- Lorenzo, Emilio (1987): “Anglicismos en la prensa”. In *Primera Reunión de Academias de la Lengua Española sobre El lenguaje y los medios de comunicación*. Madrid: Real Academia Española, 71-79.
- Luján-García, Carmen (2011): “English ‘invasion’ in Spain: An analysis of toys leaflets addressed to young children”. *English Today*, 3: 3-9.
- Luján-García, Carmen (2013): *The English Language and the Anglo-American Culture: Its Impact on Spanish Language and Society*. Newcastle upon Tyne: Cambridge Scholars Publishing.
- (n.d.) *Oxford Living Dictionaries*. OLD. Available at <https://en.oxforddictionaries.com/> [accessed on 14 November 2017].
- Pratt, Chris (1980): *El anglicismo en el español peninsular contemporáneo*. Madrid: Gredos.
- Pulcini, Virginia, Cristiano Furiassi and Félix Rodríguez González (2012): “The lexical influence of English on European languages: From words to phraseology”. In C. Furiassi, V. Pulcini and F. Rodríguez González, eds., *The Anglicization of European Lexis*. Amsterdam/Philadelphia: John Benjamins Publishing, 1-26.
- Real Academia Española. 2014. DLE. *Diccionario de la lengua española* 23rd edition. Madrid: Real Academia Española. Available at <http://dle.rae.es/> [accessed 26 September 2017].
- Real Academia Española. 2005. DPD. *Diccionario Panhispánico de Dudas*. Madrid: Real Academia Española. <http://www.rae.es/recursos/diccionarios/dpd> (accessed 24 October 2017).
- Rodríguez Díaz, José Antonio (2011): *Anglicismos y germanismos en el lenguaje de la publicidad del español peninsular contemporáneo 1998-2007*. PhD dissertation, University of Santiago de Compostela.
- Rodríguez González, Félix (1996): “Functions of Anglicisms in contemporary Spanish”. *Cahiers de Lexicologie. Revue Internationale de Lexicologie et Lexicographie*, 68(1): 107-128.
- Rodríguez González, Félix (2011): *Diccionario del sexo y el erotismo*. Madrid: Alianza Editorial.
- Rodríguez González, Félix (2012): “Anglicismos en el mundo del deporte: variación lingüística y sociolingüística.” *Boletín de la Real Academia Española (BRAE)*CXII (XXXVI): 261-285.
- Rodríguez González, Félix (2013): “Pseudoanglicismos en español actual. Revisión crítica y tratamiento lexicográfico”. *Revista Española de Lingüística (RSEL)*, 43(1): 123-170.

- Rodríguez González, Félix (2017): *Gran diccionario de Anglicismos*. Madrid: Arco Libros S.L.
- Rodríguez Medina, M^a Jesús (2016^a): “An approach to the study of the use of English in the Activities of Spanish Gyms”. *Spanish in Context*, 13(1): 128-148.
- Rodríguez Medina, M^a Jesús (2016b): “The use of Anglicisms in Spanish television Commercials of cosmetics, hygiene and personal care products”. *Hermes Journal of Language and Communication in Business*, 55: 157-169.
- Tagliamonte, Sali and Derek Denis (2008): “Linguistic ruin? Lol! Instant messaging and Teen Language”. *American Speech*, 83(1): 3-34.
- Tagliamonte, Sali (2016): “So sick or so cool? The language of youth on the Internet”. *Language in Society*, 45(1): 1-32.
- Zhang, Wei (2015): “Multilingual creativity on China's Internet”. *World Englishes*, 34(2): 231-246.

Appendix

Pure Anglicisms	Freq.	Topics	Pure Anglicisms in context	Spanish equivalents
<i>After party</i>	1	leisure	Belleza <i>after party</i> : los mejores trucos para disimular la resaca (28/11/2016)	Evento en el que personas que han estado en una fiesta se sientan a charlar relajadamente
<i>Baking</i>	2	beauty	<i>Baking</i> , el nuevo método de maquillaje (que afortunadamente no implica meter la cabeza en el horno) (21/07/2015)	Método de maquillaje
<i>Beauty</i>	3	beauty	Trenzas corsé, la nueva obsesión <i>beauty</i> (04/05/2017)	Belleza
<i>Beauty youtubers</i>	1	beauty	Algunas <i>beauty youtubers</i> han visto aumentada su popularidad gracias a tutoriales en los que explican paso a paso cómo hacerte con esta alocada tendencia. (25/02/2016)	Personas que realizan y suben vídeos a YouTube relacionados con la belleza, con frecuencia trucos de belleza
<i>Candy</i>	4	fashion	<i>Candy rooms</i> : los colores también son para el invierno (05/11/2015)	Colores vivos
<i>Celebrity / Celebrities</i>	10	leisure	Pero como todas las semanas, aunque Kylie haya sido la ganadora del título, otras <i>celebrities</i> también han conseguido estar nominadas al peor look. (14/06/2016)	Famosos
<i>Celebs</i>	1	leisure	Como todas las semanas, aunque haya habido una ganadora, hay otras <i>celebs</i> que también nos han sorprendido con sus elecciones estilísticas. (17/03/2017)	Famosos

<i>Check-out</i>	1	leisure	Desde que te reciben en recepción hasta que haces el <i>check out</i> el equipo está pendiente de que no falte nada, adelantándose incluso a tu mente (09/06/2016)	Acción de salida de un hotel o bien comprar algún producto
<i>Chokers</i>	2	fashion	Accesorios: pendientes XXL y <i>chokers</i> (29/07/2016)	Tipo de collar femenino ajustado al cuello
<i>Clown Contouring</i>	2	beauty	Después de la fiebre del contorneado que inició Kim Kardashian hace un par de años, y del nuevo método del <i>clown contouring</i> (21/07/2015)	Maquillaje femenino similar al de un payaso
<i>Cookies</i>	1	leisure	Unas <i>cookies</i> calentitas para comenzar no están nada mal. (27/01/2017)	Galletas
<i>Cottage</i>	2	fashion	Las claves para conseguir el estilo <i>cottage</i> que dinamizado con un aire casual da como resultado el conocido estilo <i>cottage</i> que nos trasladará a las más profundas campiñas inglesas (18/08/2015)	Estilo de decoración campestre
<i>Contouring</i>	1	beauty	Desde el <i>contouring</i> hasta el strobing y por supuesto los smokey eyes. (14/06/2016)	Maquillaje que realza los contornos
<i>Dot eyeliner</i>	7	beauty	¿Por qué el <i>dot eyeliner</i> es la tendencia beauty del momento? (31/01/2017)	Trazado con lápiz de ojos con puntos pequeños
<i>Dress code</i>	1	fashion	Otro elemento a tener en cuenta es saber si existe un <i>dress code</i> fijado por los organizadores del evento en cuestión y bien explicado (03/01/2017)	Estilo de vestimenta
<i>Eyeliner</i>	5	beauty	<i>Eyeliner</i> de unicornio, la magia llega a tus ojos (03/01/2017)	Lápiz de ojos
<i>Foodie</i>	2	leisure	Si vives en una gran ciudad, no tienes tiempo que perder para descubrir estos nuevos templos del arte <i>foodie</i> (02/12/2015)	Afición a la comida y a la bebida
<i>Front row</i>	1	fashion	Las modelos, diseñadores e invitadas del <i>front row</i> van de fiesta en fiesta y nosotros no perdemos detalle de sus mejores y peores looks (24/02/2017)	Primera línea
<i>Gif</i>	2	leisure	20 <i>gifs</i> de comida que te harán salivar de forma inmediata. Sigue leyendo y, sobre todo, mirando, porque estos <i>gifs</i> de comida son totalmente adictivos y generan hambre de forma inmediata. (27/01/2017)	Imágenes
<i>Girls' night out</i>	1	leisure	<i>Girls' night out!</i> cosas que suceden después de una noche de chicas (16/11/2016)	Salida de chicas por la noche
<i>Gloss</i>	3	beauty	Para unos labios llamativos y a todo color, píntalos primero de un color a modo de base y repásalos con un pincel para conseguir que esa tonalidad se vea matizada por el <i>gloss</i> (27/02/2017)	Brillo labial
<i>Glowing hair</i>	1	beauty	<i>Glowing hair</i> , la nueva tendencia capilar que te hará brillar en la oscuridad. (12/01/2016)	Pelo muy brillante
<i>Greenery</i>	2	beauty	25 imágenes que prueban que el <i>greenery</i> será el color de tu maquillaje en 2017 (15/03/2017)	De color verde
<i>Grills</i>	1	fashion	Cuando una chica normal se maquilla y arregla el pelo con la plancha para salir con sus amigas, llega el clan Jenner/Hadid y se ponen unos <i>grills</i> a medida. Se trata de unas fundas de oro que se utilizaban en los años 80	Fundas doradas para los dientes

			para decorar las dentaduras, que Madonna y Miley Cyrus recuperaron hace unos años, y que hacía tiempo que no veíamos. (24/02/2017)	
<i>Holographic lips</i>	2	beauty	<i>Holographic lips</i> , los labios también pueden brillar (27/02/2017)	Labios pintados con motivos holográficos
<i>Influencer</i>	1	beauty	Todo ha comenzado en el salón de belleza de Eun Kyung Parkm, una auténtica <i>influencer</i> en el mundo de la manicura (26/01/2017)	Persona que influye en el mundo de la moda
<i>In-flight</i>	1	beauty	<i>In-flight essentials</i> : básicos de belleza para sobrevivir a un largo viaje de avión (18/07/2016)	En un vuelo
<i>Jeans</i>	2	fashion	El look estaba formado por una cazadora vaquera con volumen en los hombros, unacamiseta de cuello alto de rejillabajo otra de manga corta y unos <i>jeans</i> de tiro alto con rayas laterales (17/03/2017)	Pantalones vaqueros
<i>Layering</i>	1	fashion	En esto consiste el <i>layering</i> , en llevar prendas de diferentes longitudes al mismo tiempo. (29/07/2016)	Estilo de vestimenta con varias capas de ropa
<i>Look</i>	47	fashion	Con esta aplicación puedes confeccionar tus <i>looks</i> de toda la semana. (26/03/2015)	Imagen, apariencia
<i>Makeup</i>	1	beauty	Ya sea a través de líneas de maquillaje con las mejores firmas de cosmética o declarando que ellas mismas se realizan sus <i>makeup</i> looks (14/06/2016)	Maquillaje
<i>Mix</i>	1	fashion	Nicki Minaj y su <i>mix</i> de tendencias, peor look de la semana (03/03/2017)	Mezcla
<i>Nail art</i>	4	beauty	En Instagram lo puedes encontrar bajo el nombre de <i>wire nail art</i> , algo así como <i>nail art</i> de alambre y consiste en utilizar un fino hilo de alambre como base de una manicura sencilla y efectiva. (26/01/2017)	Tipo de decoración de uñas
<i>Nude</i>	3	fashion	Los tonos siempre entre el <i>nude</i> y el champán (14/06/2016)	Color natural de la piel
<i>Outfit</i>	2	fashion	Natalia Vodianova a cometido un crimen estilístico esta semana. Descubre quién es el culpable de que la modelo vistiera de esta guisa y por qué nos ha defraudado tanto su outfit. (27/01/2017)	Ropa, vestimenta
<i>Over size</i>	1	fashion	La cantante eligió una chaqueta bomber rosa, una camiseta <i>over size</i> a modo de vestido y unas botas XXL. (03/03/2017)	Talla grande. Usada con el fin de llevar la ropa bastante holgada.
<i>Performance</i>	1	leisure	Tal vez la falta de libertad de entonces o de equidad en los derechos homosexuales se suplía con imaginación, carisma y <i>performance</i> descaradas (27/06/2016)	Forma de comportarse o actuar
<i>Phubbing</i>	5	leisure	¿Has oído hablar del <i>phubbing</i> ? Es la palabra que describe ese momento en el que descuidamos la compañía humana solo para echarle un ojo al móvil. Nace como una combinación de dos términos anglosajones: <i>phone</i> (teléfono) y <i>snubbing</i> (despreocupar) (22/08/2016)	Ignorar la presencia de alguien en favor de un teléfono móvil
<i>Piercing</i>	3	beauty	Uñas con <i>piercing</i> , la última manicura viral de Kim Kardashian (31/01/2017)	Perforaciones en la piel para decorar

<i>Pink lady</i>	1	beauty	Ahora vas a saber según el color de tu piel, con qué rosa lograrás destacar tus rasgos y convertirte en toda una <i>pink lady</i> (15/12/2016)	Mujer vestida y maquillada de rosa
<i>Pink lips</i>	1	beauty	¡Vivan los <i>pink lips</i> ! Elige tu rosa de labios perfecto (15/12/2016)	Labios de color rosa
<i>Pop-up stores</i>	3	leisure	Las <i>pop-up stores</i> o tiendas efímeras cada vez están más de moda y las marcas de champán se han dado cuenta. (10-10-2016)	Tiendas efímeras
<i>Rainbow Food</i>	3	leisure	<i>Rainbow Food</i> : la comida como nunca antes la habías visto. (18/05/2016)	Comida con los colores del arcoiris a base de colorantes
<i>Rainbow hair</i>	1	beauty	El <i>rainbow hair</i> ya no es solo cosa de chicas. Llega el Merman Style (26/06/2015)	Pelo decorado con los colores del arcoiris
<i>Rainbow freckless</i>	3	beauty	Llegan las <i>rainbow freckless</i> , mejillas a todo color! (25/02/2016)	Pecas de colores
<i>Reality</i>	1	leisure	Tras el éxito de su <i>reality</i> , mamá Jenner no quiso dejar pasar la ocasión y materializó la belleza de sus hijas mayores (14/06/2016)	Tipo de programa televisivo
<i>Room Service</i>	1	leisure	¿Un poco de hambre? ¡claro que sí, esto despierta el apetito! ¡ <i>Room service</i> ! Sin necesidad de levantarse, caminar con los pies descalzos sobre el frío de la cocina hasta la nevera. (20/09/2016)	Servicio de habitaciones
<i>Sane Food</i>	1	leisure	<i>Sane Food</i> - Nutrición y Emoción (30/01/2017)	Comida sana
<i>Selfie</i>	2	leisure	Su palabra preferida es <i>selfie</i> y la mayor parte del tiempo está posando (31/10/2016)	Foto tomada a uno mismo
<i>Sexy</i>	2	fashion	Gracias a su caída, la tonalidad y el escote en V, la actriz consiguió un aspecto dulce y <i>sexy</i> a partes iguales, muy a tono con su papel en el filme (03/02/2017)	Atractivo sexualmente
<i>Shopping</i>	4	leisure	Celebra el Enfemenino Day con un <i>shopping</i> apoteósico (26/10/2015)	Ir de compras
<i>Shorts</i>	1	fashion	Se llevan con faldas, <i>shorts</i> o mini vestidos fluidos. (09/02/2017)	Pantalones cortos
<i>Skyline</i>	1	leisure	Aprovechas este momento del día para salir, tomarte tu combinado favorito, charlar relajadamente y simplemente disfrutar de las preciosas vistas del <i>skyline</i> de la ciudad desde una terraza y sentirte un poquito como la reina del mundo (02/12/2015)	Horizonte
<i>Slim</i>	1	fashion	El look compuesto por un vestido corto de cuadros <i>vichy</i> , con capa del mismo estampado, pantalones de piel de corte <i>slim</i> y botas de caña alta, no eran lo más adecuado para esa tarde. (27/01/2017)	Esbelto, delgado
<i>Slip dress</i>	1	fashion	Kendall eligió un delicado <i>slip dress</i> y lo combino con una sudadera de crochet con enormes borlones (24/02/2017)	Vestido de tirantes
<i>Soft</i>	1	fashion	Existen muebles de estética " <i>soft</i> " que encajarán a la perfección tanto en líneas puras como redondeadas. (05/11/2015)	Suave

<i>Spa</i>	1	leisure	Su concepto de <i>Oenoterapia</i> y <i>Spa Sommelier</i> te permitirá probar algunos de sus vinos de la mano de un experto, analizarlos y descubrir cuál es el que más te gusta (09/06/2016)	Lugar de tratamientos acuáticos relajantes
<i>Streetstyle</i>	4	fashion	La hemos visto en la pasarela, el <i>streetstyle</i> y por supuesto en las <i>celebrities</i> . (08/02/2017)	Estilo urbano
<i>Sweet</i>	1	fashion	Muebles " <i>sweet, sweet</i> " (05/11/2015)	Dulce
<i>Tip</i>	2	fashion	Antes de seguir con los <i>tips</i> , te dejamos un <i>shopping</i> con prendas y accesorios que pueden encajar para distintos eventos formales (03/01/2017)	Truco
<i>Top</i>	1	fashion	Aún así, lo que logró llevarse toda la atención fue el cinturón que Halsey llevó a modo de <i>top</i> . ¿Estaría cómoda? Para combinarlo eligió una chaqueta larga tipo impermeable y unos pantalones. (10/03/2017)	Camiseta
<i>Trendy</i>	1	fashion	Toma nota de nuestros consejos y serás la más <i>trendy</i> de la temporada. (29/07/2016)	De moda o tendencia
<i>Tweed</i>	2	fashion	<i>Tweed</i> en el armario. El tejido favorito de Coco Chanel arrasará esta temporada (20/08/2015)	Mezcla de lana
<i>Welcome pack</i>	1	leisure	Te obsequiarán con un <i>welcome pack</i> (no queremos desvelar todas las sorpresas). (26/10/2015)	Paquete de bienvenida
<i>Wire nail</i>	1	beauty	Sin embargo cada vez son más las publicaciones web que animan a que tú misma te hagas tus diseños <i>wire nail</i> usando el típico alambre que se vende para hacer pulseras y demás accesorios. (26/01/2017)	Decoración de uñas con alambres
<i>Working girl</i>	1	fashion	Encontrarás los looks <i>working girl</i> más buscados. (22/11/2016)	Estilo para ir al trabajo

Table 4: Non-adapted or pure Anglicisms compiled from *enfemenino*

Adapted Anglicisms	Freq.	Topics	Adapted Anglicisms in context	Spanish equivalent
<i>Climax</i>	1	leisure	Utiliza alguna de estas canciones y el <i>climax</i> está asegurado (20/09/2016)	Apogeo, culminación
<i>Tartán</i>	1	fashion	Tartán pero en <i>total look</i> (20/08/2015)	Tipo de tejido con cuadrados tipo escocés

Table 5: Adapted Anglicisms compiled from *enfemenino*

False Anglicisms	Freq.	Topic	False Anglicisms in context	Spanish equivalents	Real English equivalent
<i>Afterwork</i>	11	leisure	Con un ambiente relajado y una conversación agradable no nos cabe duda de que el <i>afterwork</i> es el impulso que necesitas para continuar con ánimo la semana. (02/12/2015)	Cóctel o copa que se toma con los compañeros de trabajo tras una jornada laboral	Afterwork party

<i>Bómbler</i>	3	fashion	Por si no te has enterado, la <i>bómbler</i> es la chaqueta del momento (08/02/2017)	Tipo de chaqueta	Bomber jacket
<i>Bombers</i>	1	fashion	Calcetines de rugby, <i>bombers</i> , sweaters y abrigos acolchados te vestirán de arriba abajo. (29/07/2016)	Tipo de chaqueta	Bomber jacket
<i>Fashion</i>	3	fashion	Natalia Vodianova, apodada <i>Supernova</i> , es una de las supermodelos más conocidos del mundo, su exitosa carrera y pasión por la moda le han convertido en un personaje clave del panorama <i>fashion</i> al que nunca le quitamos el ojo (27/01/2017)	Moda	Fashionable
<i>Reality</i>	1	leisure	Tras el éxito de su <i>reality</i> , mamá Jenner no quiso dejar pasar la ocasión y materializó la belleza de sus hijas mayores (14/06/2016)	Tipo de programa televisivo	Reality show
<i>Short</i>	1	fashion	No podría faltar el ya mencionado <i>short</i> en satén fucsia, la camisa estampada, la cazadora de ante con mangas de pelo... (28/10/2016)	Pantalón corto	Shorts
<i>Sport</i>	1	fashion	Eso sí, hay que tener en cuenta el diseño porque si es muy <i>sport</i> serán más adecuadas para looks informales. (09/02/2017)	Estilo deportivo	Sporty
<i>Teenager</i>	1	fashion	Todo muy <i>teenager</i> , para tener contento a su público. (14/06/2016)	Adolescente	Teenage

Table 6: False Anglicisms compiled from *enfemenino*

Hybrid Anglicisms	Freq.	Topics	Hybrid Anglicisms in context	Spanish equivalents
<i>Country chic</i>	1	fashion	Piezas chics que construyen los detalles de este estilo evocando los ambientes más románticos y bucólicos acercándose incluso al <i>county chic</i> (18/08/2015)	Estilo de ropa
<i>Ombre lips</i>	3	beauty	<i>Ombre lips</i> : pásate al efecto degradado en tus labios (19/06/2016)	Labios maquillados con pintura degradada

Table 7: Hybrid Anglicisms compiled from *enfemenino*

**Relevance Theory Online Bibliographic Service
(AUTHORS' INDEX)**

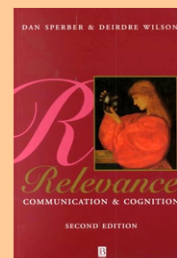
Last modified: **1 March 2016**

Created and updated by

Francisco Yus

University of Alicante (Spain)

[<https://sites.google.com/site/franciscoyus>]



**RELEVANCE THEORY ONLINE
BIBLIOGRAPHIC SERVICE
(last updated 1 March 2016)**

Thematic Sections:

**[http://personal.ua.es/francisco.yus/
rt.html](http://personal.ua.es/francisco.yus/rt.html)**

Authors' Index:

**[http://personal.ua.es/francisco.yus/
rt2.html](http://personal.ua.es/francisco.yus/rt2.html)**

The influence of the English language on the description of cosmetic products

Cristina Tejedor Martínez
Universidad de Alcalá de Henares
cristina.tejedorm@uah.es

ABSTRACT

The influence of English as an international language has increased in many areas, from scientific, technological, economic and political fields to cinema, music and advertising. The use of English as a global language has exerted an enormous pressure over languages, especially on the lexical level. Since the second half of the 20th century, many works have already dealt with Anglicisms in Spanish (most cases of recent borrowings) in various fields, including fashion (Balteiro and Campos, 2012) and television cosmetics commercials (Rodríguez Medina, 2016a), but the study of descriptions provided by the brands for their cosmetic products has received less attention in Spanish. This paper provides an analysis of facial cosmetics descriptions selected from a corpus collected in 2016 from four Spanish cosmetic brands. Language creativity exploiting both the use of English borrowings and the influence of the English language in some orthographical patterns related to word-formation processes in Spanish in this genre will be discussed. The proportion of the influence of the English language on this kind of texts may be an important factor in determining its socio-psychological effect on the target public; besides, the quantitative results will be compared with those obtained in our previous studies in the fields of tourism and computing. A qualitative analysis of a selection of examples from our corpus will be offered. The present study intends to illustrate the influence of the English language on the information consumers can read about cosmetic products.

Keywords: Language of cosmetics, Anglicisms, false Anglicisms, English influence on orthographical word-formation patterns



1. Introduction

The study and research of the Cosmetology field has increased in the last years due to the socio-economic impact of the cosmetic industry on the world economy. The use of cosmetics is not a new practice; in fact, all civilizations have used cosmetics for different purposes, such as in religious rituals, to enhance beauty, and to promote good health. The history of cosmetics usage is as long as humankind; in fact, “early tomb paintings, frescoes, and mosaics all suggest that the use of cosmetics was widespread among people in ancient Mesopotamia and Egypt” (*The Use of Cosmetics in Bible Times*). The search and importance of beauty is a natural tendency in human beings and has been sought since ancient times in an attempt to change people’s appearance and benefit from being attractive.

Cosmetology is defined as “the cosmetic treatment of the skin, hair, and nails” (*Merriam Webster*); “the professional skill or practice of beautifying the face, hair, and skin” (*Oxford English Dictionary*); or “the art or profession of applying cosmetics” (*The Free Dictionary*). The term *cosmetics* derives from the Greek adjective *κοσμητικοξ*, which means ‘related to adornment’, and is defined as “of, relating to, or making for beauty especially of the complexion” (*Merriam Webster*); “a preparation applied to the body, especially the face, to improve its appearance” (*Oxford English Dictionary*); or “a preparation, such as powder or a skincream, designed to beautify the body by direct application” (*The Free Dictionary*).

Cosmetic products have managed to have great impact on our lives, fashion, culture and history. The reason for the popularity of cosmetic products is the important physiological and psychological benefit they impart to the consumer. Several studies have shown that, apart from the physiological effects, “there are very real psychological and social benefits to be gained from the use of cosmetics. It is these benefits that ensure cosmetic products are, and will remain, an integral part of life” (King, 2013). Therefore, consumers use cosmetic and personal care products every day to protect their health, ensuring their well-being and boosting their self-esteem, because they endeavour to look good and younger, also to be healthy, but enhancing beauty. Cosmetics play an important role in boosting one’s inherent beauty and physical features, and skin products are one of the most widely used by individuals. There is a progressive demand of cosmetic products that has led to the growth of the cosmetic market across the world. The importance of the cosmetics industry is growing in global economies and the distribution is part of the international economy. In fact, the cosmetic sector has been rising significantly over the last few years due to the increase of disposable income and the widespread concern about beauty. Several Spanish brands of cosmetics happen to be appreciated in the global cosmetic market. Therefore, four Spanish cosmetic brands internationally positioned in the market were selected to carry out this study on how English language affects the Spanish lexicon.

2. Theoretical framework

The English language has spread vastly, as the modern *lingua franca*, increasing its influence on the global market over the last few years, affecting every domain of the linguistic system. As pointed out by Kachru (1994: 135), “it is for the first time in linguistic history that a language has established contact with practically every language family, both formally and functionally”. The importance of English, its global dominance in science, technology, commerce, marketing, advertising, as well as sports, along with the concepts of prestige and modernity attached to it, have reinforced the power of this language around the world and have facilitated the continuous incorporation of English borrowings in Spanish in the last decades in a number of specialised fields.

The influence of English upon the Spanish language can be traced back to the Renaissance period, but the massive influx of borrowings is documented over the past sixty years. The influence of the English language on many different technical and scientific domains is undeniable, not only in Spanish but in most European languages (Görlach, 2001, 2002). The process of “Anglicization of European languages”, as some authors have denominated it (Pulcini, Furiassi and Rodríguez González, 2012: 3), has been possible because of the constant exchanges between English and the other languages since the 18th century. Although it is noticeable in all levels of language, its influence is most salient in the lexical and semantic fields.

Ever since the 1960s, several scholars have studied the phenomenon of Anglicisms, paying attention to theoretical aspects and analysing how these borrowings are being used and adapted into the recipient system. In fact, as Rodríguez Medina (2016b: 128) states, “[t]he study of Anglicisms in Spanish as a complex result of language contact and cultural globalization has increasingly caught scholars’ attention”. The pioneer in this field was Lorenzo Criado, who published several studies on the topic in the 1950s and continued doing so until the end of the 20th century (1996; 1999). Pratt (1980) shared many of Lorenzo’s views and like him concentrated mainly on lexical Anglicisms, establishing a categorization of English borrowings and analysing the orthographic, morphological and syntactic influence on the recipient language.

Another important scholar in the study of Anglicisms in the Spanish language is Rodríguez González who has covered diverse issues that had to do with borrowings: the process of semantic and morphological translations in the coining of calques (2002b, 2004) and analysed the spread of Anglicism in various jargons, like the language of sports (2012). He published in collaboration with Lillo Buades (1997) the most important dictionary on the topic, after the one published by the Panamanian linguist, Alfaro (1970). Rodríguez González also collaborated in the *Dictionary of European Anglicisms* (2001) edited by Görlach; after these lexicographical works, he explained the criteria that should be taken into account for the compilation of a corpus (2003) in order to elaborate this type of dictionaries and he has just published a new work, *Gran diccionario de anglicismos* (2017).

Many others contributed to the discussion of Anglicism, like Gimeno and Gimeno (2003) with a comprehensive view of the phenomenon within the languages in contact

framework established by Haugen (1950) or Weinreich (1953). Gómez Capuz (1998, 2000, 2004) worked on the classification of borrowings and paid attention to the group of pseudoanglicisms. Oncis Martínez (2012) approached Anglicisms in Spanish, mainly calques and semantic borrowings, by means of a corpus-based analysis, searching the CORDE (*Corpus Diacrónico del Español*) and the CREA (*Corpus de Referencia del Español Actual*) corpus, as he considers that their use “have become an indispensable tool for research on English loanwords” (2009: 116). These and other linguists, who have published on the topic, have helped to shape our conception of Anglicism and our way of dealing with the borrowing process.

Regarding the field of cosmetology, several authors have paid attention to the study of borrowings in Spanish in the field of fashion, which includes cosmetic products. Balteiro and Campos (2012) concentrated on the analysis of false Anglicisms in the Spanish language used in fashion and beauty, and compiled a corpus searching in lexicographic sources and online webpages to check uses of the examples. Their work concluded that “false anglicisms are an endless source of information on morphological and semantic evolutions in language contact [...], as part of the Anglicization of European lexis” (2012: 250). Rodríguez Medina (2016a) analyzed the use of Anglicisms in Spanish television commercials of cosmetics, hygiene and personal care products, and explained the reasons why Anglicisms are frequent in Spanish advertising, coming to the conclusion that their presence is related to the prestige of the products (2016a: 168).

Finally, it is also important to mention the previous work by Rodríguez Segura (1999) concerning the study of Anglicisms in the mass media. Her work included a list of Anglicisms, classified and illustrated with contextual information and explanations, and pointed out that this linguistic and sociological phenomenon was becoming important and more complex, as can be assured almost two decades after her study.

Descriptions of the cosmetic products are closely related to their advertising and commercialization processes. Several researchers (Bhatia, 2001 and 2006; Hsu, 2008; Martin, 2002, among others) have shown that English code-switching in advertising is a universal phenomenon as English language predominates in publicity and marketing. Language contact situations, the power of mass media, international marketing and advertising, information and communication technologies and, very often, poor and superficial translations have been regarded as the reasons for the widespread phenomenon of Anglicization.

3. Methodology

This study is included in the work carried out by the research team aLiLex (Acronym of the Spanish equivalent to *Lexical Linguistic Analysis — Análisis Lingüístico del Léxico*) starting back in 2003 and 2004. The work began dealing with the process of a textual corpus creation working firstly on texts related to computer science. In the following years, we based our research on the compilation and analysis of various genres from the tourist field examining several sources. Thus, we covered the healthcare world with specialized publications aimed at physicians, but also with more public outreach

magazines; and, finally, magazines aimed at a general audience interested in topics related to science and technology. In addition, we compiled documents and information on business and economy, sports, politics, general language in the mass media, etc. Some authors have been using corpora for the study of Anglicisms (Oncins Martínez, 2009; Balteiro Fernández, 2011). In fact, nowadays researchers consider that “in the study of Anglicisms, corpora are indispensable because they offer up-to-date source material from Anglicisms or new meanings/senses of Anglicisms may be detected” (Pulcini, Furiassi and Rodríguez, 2012: 18).

For the present article, we concentrate on the specialized field of cosmetology, so a corpus to study the presence of borrowings and the influence of English on this field has been compiled. As in previous research, we use authentic sources from the specialized domain. In this case, four Spanish commercial brands of cosmetics were selected: *Germaine de Capuccini* (Alicante), *Natura Bissé* (Barcelona), *SKinClinic* (Alicante) and *Farma Dorsch* (Madrid). Although all of them have international expansion, their trade volume differs and the number of lines and cosmetic products in each one varies too. The description of facial products, all the lines launched by those brands in 2016, form part of this study. The leaflets that accompany the products, also published in the web pages, were collected in order to prepare the corpus. The leaflets are divided in several sections, as cosmetic products must include information that explains what they are for, how to use them safely, and how to obtain the best result;¹ the main sections are the characteristics of the product, the ingredients, and the indications of how to use it. For this study, the ingredient lists and the name of the products have not been analyzed, although some ingredients, and even sometimes the names, appear in the data because they are mentioned in other sections of the leaflet.

Once the corpus was compiled and divided in four sub-corpora, SimpleExtractor² tool was used to extract terms, in order to obtain a list of candidate terms and the information about their frequency and their contexts of occurrence. After refining the list of candidate terms, the final list of Anglicisms³ to analyze was obtained and were added to our database, named *anglicor*.⁴ This database stores linguistic data on English loanwords. The database includes: grammatical information, such as category, gender and number; semantic information, such as definition and semantic field; etymology; graphic marking of loanwords; lexicographical information; and socio-pragmatic information. In this first step of the study, only parts of these fields were completed in the database for each Anglicism recorded. In figure 1, an example of the database is provided.

The screenshot shows the 'NUEVA UNIDAD LÉXICA' (New Lexical Unit) form in the ANGLICOR system. The interface includes a header with the ANGLICOR logo and user information (USUARIO: Isabel Cruz). Navigation buttons for 'LISTAR TODOS', 'NUEVA PALABRA', and 'BÚSQUEDA' are visible. The main form is divided into several sections:

- *Unidad léxica:** A text input field for the word.
- Información Gramatical:**
 - Categoría:** A dropdown menu set to 'Seleccionar'.
 - *Género:** A dropdown menu.
 - *Número:** A dropdown menu.
 - Adaptación:** Checkboxes for 'Ortográfica', 'Morfológica', 'Flexiva', and 'Derivativa'.
 - Pronunciación:** A text input field with a small icon for pronunciation.
- Información semántica:**
 - *Definición:** A large text area for the definition.
 - *Campo semántico:** A dropdown menu set to 'Seleccionar'.
- Etimología:**
 - OED Inglés:** A text input field.
 - Otra Lengua:** A text input field.
 - Americano Inglés:** A text input field.
 - Otra Lengua:** A text input field.
 - Etymonline Inglés:** A text input field.
 - Otra Lengua:** A text input field.
 - DRAE Inglés:** A text input field.

At the bottom right, there is a 'Cambiar nivel de zoom' button.

Figure 1: Anglicor record

Although the classifications of borrowings from some of the previous authors mentioned above have been consulted, we have adapted the criteria to classify English borrowings in two groups, in order to compare the results with those previous studies that included this field:

- Anglicisms, words that come directly from English or English was the language of transmission, with no adaptation or partially adapted to Spanish, or derivative words, whereby an element from English origin is combined with a Spanish morpheme. Consequently, we have grouped several types of borrowings under this general heading.
- False anglicisms, also named pseudoanglicisms. Gómez Capuz (2004: 63-67) dealt with the different kinds of pseudoanglicisms or false borrowings. Recently, Furiassi (2010), among other researchers, studied these cases of borrowings and pointed out the problems of their identification. Balteiro and Campos (2012: 234-236) provided a review of the literature on the description and research on false Anglicisms.

All lexical items found in the corpus were looked up in several lexicographic sources:⁵ general Spanish monolingual dictionaries, such as *Diccionario de la Lengua Española* (DRAE), *Diccionario Panhispánico de dudas* (DPD),⁶ and works dealing with Anglicisms, such as the *Nuevo diccionario de Anglicismos* (NDAng), *A Dictionary of European Anglicisms* (DEAng) and *Gran diccionario de anglicismos* (GDAng)⁷. Besides, some items have also been checked in English monolingual dictionaries, such as the *Oxford English Dictionary* (OED), *Oxford Living Dictionaries* (OLD) and *Longman Dictionary of Contemporary English* (LDCE), regarding their meaning and use in English. We have decided to consult OLD and LDCE as both dictionaries use authentic

corpora in their compilation, therefore they record present-day English language which takes some time to be incorporated into other lexicographical works like the OED. The present study intends to illustrate the influence of English language on the information consumers can read about cosmetic products, paying special attention to lexical Anglicisms.

4. Results

4.1. Quantitative and qualitative analysis of Anglicisms

Although the main aim of our study is the qualitative analysis of the results obtained, firstly, we would like to offer some quantitative information from our research and will contrast the results with our previous studies dealing with the introduction of English loanwords in Spanish. Secondly, we will discuss several examples of the Anglicisms found in our study, comparing them with previous related research in the field.

The corpus compiled has 47,089 tokens in total, subdivided as follows in the four sub-corpora:

- Germaine de Capuccini: 13,726 tokens
- Natura Bissé: 17,899 tokens
- SkinClinic: 9,406 tokens
- Farma Dorsch: 6,058 tokens

From the analysis of the corpus, a total number of 51 different types of Anglicisms have been found. These lexical types distribute in the four sub-corpora as follows:

	Germaine de Capuccini	Natura Bissé	SkinClinic	Farma Dorsch
Number of lexical types	32	24	14	13

Table 1: Number of lexical types

The distribution of these lexical types in the four sub-corpora shows that 23 types are recorded in at least two of the cosmetic brands and 28 lexical types appear in one of the brands.

Some of the data related to the cosmetic brands analyzed are shown in the following table:

Brand	N° of tokens	N° of types	% type/token
Germaine de Capuccini	13,726	148	1.07%
Natura Bissé	17,899	122	0.68%
SkinClinic	9,406	106	1.12%
Farma Dorsch	6,058	65	1.07%
TOTAL	47,089	441	0.93%

Table 2: Type-token ratio: cosmetic brands

If we compare the results in the four brands, regarding the number of types in relation to the number of tokens, the brand SkinClinic has the highest type-token ratio. If we analyze the different types of Anglicisms recorded in relation to the types in each sub-corpus, the brands Germaine de Capuccini and Natura Bissé offer a richer variety; however, in the brand SkinClinic, Anglicisms are repeated very often in the descriptions. We can relate these results to the international growth of both brands, Germaine de Capuccini and Natura Bissé. In fact, both are very successful in the Asian cosmetics market, as Spanish cosmetics are gradually gaining their presence in China; for example, Germaine de Capuccini won the prize “Marca Líder en el Sector Profesional” in China Beauty Expo 2014.

As this study would be integrated in the corpus of specialized areas that is being compiled, we have attempted to check whether the presence of Anglicisms can be considered significant through frequency data in the three areas studied. If we compare the analysis of the impact of Anglicisms in two specialized languages previously analyzed (De la Cruz and Tejedor, 2012), Tourism and Computer, with the field of Cosmetology,⁸ we can conclude that the number of English loanwords in Spanish computer language is obviously higher, similarly to what we deduced when comparing the fields of tourism and computer.

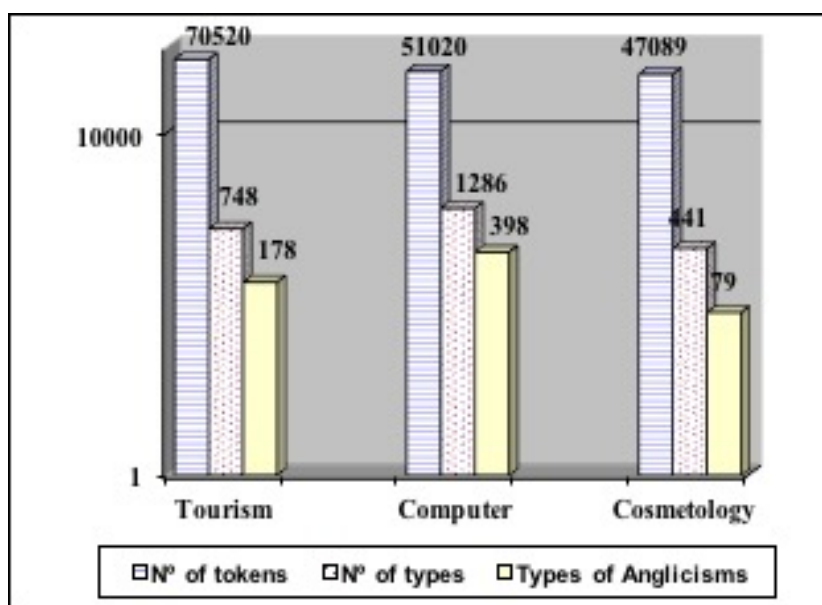


Figure 2: Presence of Anglicisms in three specialised languages

The use of technology has grown exponentially worldwide and it is contributing to the gradual expansion of English; moreover, this specialized area provides more borrowings from English to other languages. In addition, loanwords adopted from the Computer language affect not only general language, but also other specialized fields. For example, the Anglicism *3D* is used in the description of technology developed for skin care creams. Besides, borrowings successfully transferred across different specialized languages. Some examples found in our corpus demonstrate that the field of Cosmetology also incorporates lexical items from the field of photography, as the cases of *flash* having expanded its meaning as a beauty treatment that “aporta un rápido efecto y a la vez es de efímera duración, que tiene por objeto lograr resultados estéticos inmediatos tales como una piel descansada, más luminosa, que luzca recuperada y con expresión de relajación, eliminando los típicos signos de fatiga”,⁹ and *soft focus*, as the “optical blurring of skin wrinkles caused by manipulating the transmission and scattering of light from and into skin”,¹⁰ or lexical items from the field of psychology, as *stress-reduction* “to reverse the effects of stress-related skin conditions”.¹¹ All research carried out in different specialised fields during the last decades indicates that English loanwords are increasing their presence in the Spanish lexicon. Although quantitative analysis offers helpful information and points out the enormous amount of Anglicisms that have been adopted in the last decades, the qualitative study puts forward new factors to better understand other aspects of the borrowing process and how it affects the Spanish language.

From the 51 different lexical types of Anglicisms, 11 are recorded by the *Diccionario de la Real Academia Española* and two more are included in the *Diccionario Panhispánico de Dudas* (2005). Rodríguez González and Lillo Buades (1997) registered 18 Anglicisms and three new ones have been included in the new dictionary of Anglicisms (Rodríguez González, 2017), i.e. *anti-aging*, *detox* and *serum*. However, in

three cases in the first dictionary and four cases in the second one, the meaning recorded by the dictionaries does not correspond with the sense of the Anglicisms in our corpus. Regarding Görlach's lexicographical work (2001), 21 of the lexical types found in our corpus are included in the dictionary, but only 17 are recorded as being used in the Spanish language and four cases are not registered with the meaning of the Anglicisms in our corpus, i.e. *booster*, *coach*, *detox* and *flash*.

According to the criteria established for the classifications of the items, we have detected the following 18 false Anglicisms, representing a total of 35,29%, and 33 Anglicisms, standing for 64,70% in the corpus.

False Anglicisms: *acti-aging*, *anti-estrés*, *airless*, *booster*, *coach*, *flash*, *hydractivas*, *hydracure*, *hydra-flow*, *inflamm-aging*, *lifting*, *no-stress*, *peeling*, *phyto-complejo*, *relax*¹², *stress-reduction*, *soft focus* and *water-oil-free*.

Anglicisms: *after-shave*, *aftersun*, *Aminoessence Cocktail*, *anti-aging*, *cocktail*, *confort/comfort*, *detox-diet complex*, *detox*, *detoxificar*, *detoxificación*, *detoxificante*, *detoxificador*, *estrés*, *eyeliner*, *film*, *ginger*, *kit*, *laser/láser*, *lipo-filling*, *oil-free*, *packaging*, *parabenes/parabenos*, *resilencia*, *roll-on*, *scrub*, *serum/sérum*, *3D*, *test*, *testado*, *voluminizador*, *voluminizar* and *waterproof*.

If we compare our results with the study about Anglicisms in TV commercials (Rodríguez Medina, 2016a: 162), from the 43 Anglicisms she recorded in the section 'anti-aging and sun creams', 35 were pure Anglicisms (81,3%) and only 8 pseudo-Anglicisms (18,6%). Therefore, the percentage of false Anglicisms is higher in our corpus, even though we have not analyzed the ingredients and names of the products.

The documentation process for the identification of false Anglicisms is quite complex as there are some lexicographical sources that record them as if they were Anglicisms, but when we compare the meanings and uses in monolingual dictionaries, we check that these items do not exist or they do not have the attributed meaning in the original language. We agree with Rodríguez Medina (2016a: 162-163) when she pointed out that pseudo-Anglicisms "are very useful to create English-looking pseudo-scientific words to name products" and not only for the name of the products, but also for the characteristics that they may have when being used by the consumers. In fact, sometimes one brand coins a new item or uses an existing one widening its meaning and becomes popular (well accepted by consumers), other brands incorporate it in future promotions or information of their products or try to create a similar one that can be associated with this brand or this product. Besides, we have also intended, while documenting the cases of false Anglicisms, to check the possibility of re-borrowing, "a false Anglicism may be re-imported into English through usage in other languages" (Balteiro and Campos, 2012: 238), by observing their occurrences in web pages. This re-borrowing process is quite complex as even the documentation process can be misleading.

Regarding the Anglicisms found in the corpus, only seven terms have been recorded in the four cosmetic brands: (efecto) *flash*, *laser/láser*, *lifting*, *parabenes/parabenos*, *peeling*, *sérum* and *test*; and only ten terms are included in at least two of the sub-corpora: *after-shave/aftershave*, *after-sun/aftersun*, *confort/comfort*¹³, *estrés*, *detox*, *detoxificante*, *film*, *kit*, *oil-free* and *testado*. We will comment the four most interesting cases in each of

these two groups (see appendix B for contexts and information from lexicographical sources and web pages consulted):

- *Flash*: This word is recorded as an Anglicism in the four main lexicographical works consulted (DRAE, NDAng, DEAng, GDAng), but its meaning refers to photography in all cases; therefore, a semantic extension or semantic shift has occurred. The change in meaning indicates that it should be classified as a false Anglicism in the specialised area of Cosmetology, meaning a beauty treatment that

aporta un rápido efecto y a la vez es de efímera duración. Tiene por objeto lograr resultados estéticos inmediatos tales como una piel descansada, más luminosa, que luzca recuperada y con expresión de relajación, eliminando los típicos signos de fatiga como pueden ser la sequedad o la falta de luminosidad y tonicidad.¹⁴

Whether the semantic extension has occurred spontaneously in Spanish or through the influence of English, it is not totally clear as the term *flash* is also being used nowadays in English for facial creams with softening and brightening results in the skin.

- *Lifting* (corresponding to the English *face-lift*) and *peeling* (corresponding to the English *exfoliation* and *facial scrub*): Although both are included in the three specialized works on Anglicisms (NDAng, DEAng, GDAng), *peeling* is not recorded in DRAE, whereas *lifting* is. For the classification of the first term, we agree with Balteiro and Campos (2012: 247) who explained that the first one is a false Anglicism, because it exists in English but it is used in Spanish with a different meaning and “it is not clear whether the metaphorical extension has occurred spontaneously in Spanish or through the influence of English”. In fact, the term *facelift* means “a cosmetic surgical operation to remove unwanted wrinkles by tightening the skin of the face” (OED);¹⁵ the second element of the compound is used in Spanish to refer not only to the surgical operation, but to the effects some facial products would have in the skin, the action of lifting it and helping diminish visible signs of aging. In the case of the term *peeling*, the English verb *peel* is used in Spanish to refer to the “thick substance which you use to clean the skin on your face thoroughly” (LDCE) and to “wash or rub (a part of the body) with a granular substance to remove dead skin cells” (OLD). The word *peeling* is used in the English compound *peeling skin* (desquamation) which means “damage and loss or shedding of the outer layer of your skin (epidermis)”;¹⁶ also, the word *peelings* does exist in English, meaning “strips of the outer skin of a vegetable or fruit” (OLD). Nevertheless, it is difficult to state whether the Spanish loanword has been coined by the ellipsis of one element in the compound *peeling skin* and a semantic shift, or a metaphorical extension of *peelings* to a different meaning in Spanish, “exfoliación suave de la superficie

- epidérmica” and “producto empleado en el tratamiento” (GDAng). However, we consider that it is possible to conclude that this loanword is also an example of false Anglicism.
- *Parabenes/parabenos*: DRAE has included this Anglicism and, although both spellings have been found in our corpus, it recommends the use of *parabeno*, but the use of both spellings is extended, becoming more common the one recommended by Real Academia de la Lengua.
 - *Serum/sérum*: This item is only recorded in Rodríguez González (2017: 884), being defined as “cosmético con textura en forma de gel que se aplica a la cara para hidratar, tensar y dar brillo a la piel”. He also explains that “El término es bien acogido en el campo de la cosmética en detrimento de su equivalente español *suero*, tan ligado al ámbito de la medicina” (2017: 884). The recommended term *suero* is only used by Farma Dorsch. Nevertheless, this meaning has not been found in any of the lexicographical English works consulted as part of this study. After expanding our search, two different English sources record this item: the *Merriam Webster Dictionary* defines serum as “a usually lightweight cosmetic preparation especially for use on the face; specifically: a typically water-based, often concentrated preparation that lacks lubricating and thickening agents”; and an article about skin care products provides information about the use of the product since mid 1990s and how the term comes from professional cosmetology.¹⁷As a consequence, this term should be considered an Anglicism.
 - *Aftershave* and *aftersun*: Both words are included in the three lexicographical reference works specialised in Anglicisms, indicating that their use in English is *aftershave lotion* and *aftersun lotion*, but they are not recorded in DRAE. Accordingly, they should be classified as false Anglicisms and this is the case in the study of Rodríguez Medina (2016a: 166). However, after checking both lexical items in the OED, we have confirmed that they do exist in the English language with the same meanings. *Aftershave* is defined as “any of various preparations and products applied to the skin after shaving; (now esp.) an astringent scented lotion applied in this way” and *aftersun* as “a product applied to the skin after exposure to the sun to help reduce sunburn, rehydrate the skin, or enhance tanning”. Besides, both compounds appeared spelled with space, and with/without hyphen. Consequently, we decided to include them in the group of Anglicisms.
 - *Detox*: Similarly to the example of *serum*, it is only recorded in Rodríguez González (2017: 287-288) with the meaning “referido a un proceso de desintoxicación o purificación, especialmente a la eliminación de toxinas mediante zumos. El método se ha popularizado hoy entre los que cuidan de su salud por influencia de nutricionistas y dietistas”. Even though this meaning is not recorded in the OED, a semantic widening has taken place in the English language before the borrowing process. Some dictionaries do show the new meaning of this word: “when you do not eat solid food or only drink special liquids for a period of time, which is thought to remove harmful substances from

your body” (LDCE). In conclusion, after checking all the sources, we decided to include it in the group of Anglicisms as the term is also used in English to refer to creams that detoxify the skin.

- *Oil-free*: This word is not included in any of the lexicographical sources, not even in the OED, although many English web pages on cosmetics use this compound. Again, we decided to check a monolingual dictionary compiled on corpus-based data and the lexical item is recorded with the meaning: “an oil-free liquid, skin treatment etc. contains no oil” (LDCE). Hence, as in the previous example, we classify it as an Anglicism.

Other Anglicisms found in the corpus that we consider worth commenting in this study are:

- *Anti-aging*: In this case, we decided to change the previous classification as pseudo-Anglicism (Rodríguez Medina, 2016a: 166), due to new meaning of the word recorded in the OLD: “(of a product or technique) designed to prevent the appearance of getting older”. Also, it has been included in Rodríguez González (2017: 23-24) as “referido a los tratamientos y productos que tienen la propiedad de evitar o retrasar el envejecimiento de la piel”. As a result, it was included in the group of Anglicisms.
- *Resilencia*: it is recorded in DRAE as Anglicism “Del ingl. Resilience” and defined as “capacidad de adaptación de un ser vivo frente a un agente perturbador o un estado o situación adversos”. The word *resilience* is defined in OED as “elasticity; the power of resuming an original shape or position after compression, bending, etc.”. Nevertheless, we found it complicated to classify this item, as it seems that the meaning recorded in both English and Spanish lexicographical sources refers to other semantic fields, not to Cosmetology, but the word is used in both languages to refer to the skin capacity to regenerate and fight external factors, improving its elasticity and youthful appearance. Whether the semantic extension occurred in the English language first, being then adopted in the Spanish language, it is difficult to state, although this could be the most probable borrowing process.
- *Airless*: The term *airless* is used as it were an English term for ‘sistema al vacío’ in our corpus, but this is not the real meaning of the English word, as can be checked from the definitions in the OED¹⁸ and OLD.¹⁹ It is not included in DRAE nor in any of the specialized dictionaries consulted. The compound *airless injection* is recorded in English and means “Injection of fuel into the cylinder of a diesel engine by means of a fuel pump rather than by compressed air; frequently attributive” (OLD). After searching the web, we have found the term *airless* used for a type of paint sprayers²⁰ and for bottles and containers used for cosmetic products;²¹ therefore, the term exists in the English lexicon, but the new sense has not been recorded in dictionaries.

4.2. Analysis of the influence of English word-formation orthographical patterns on Spanish

The influence of the English language does not only affect the borrowing process, but also appears in the orthographical patterns of Spanish derivatives and neoclassical compounds. We will analyze the list of prefixed items and neoclassical compounds found in our corpus (see table 3). In fact, the use of some prefixes and combining forms in Spanish words followed by a hyphen imitates the English orthographical structure. The prefixes and combining forms used in the examples in table 3 do exist in Spanish, according to RAE, but the orthographical pattern does not follow the rules in the Spanish language. According to RAE (2010: 535), “no se consideran ortográficamente adecuadas las grafías en las que el prefijo aparece unido con guion a la palabra base (*anti-mafia, *anti-cancerígeno) o separado por ella por un espacio en blanco (anti mafia, anti cancerígeno)”. Similarly, RAE explains that the combining forms used in these formations “Si va antepuesto, se denomina elemento compositivo prefijo: biodiversidad, ecosistema; si va pospuesto, se denomina elemento compositivo sufijo: antropófago, neuralgia” (DPD) and the examples provided in the explanation clearly show that the hyphen is not used in the Spanish word-formation process, for example, *biodegradable* is recorded in RAE, but without the hyphen. Similarly, Fundeu²² (2017) explains that “el prefijo **anti-**, como todos los prefijos, se escribe por regla general **unido a la palabra a la que precede sin guion**” and that “el elemento compositivo eco-, que significa entre otras cosas ‘ecología’ y se emplea en términos como ecoparque, ecotasa o ecoterrorismo, se escribe unido a la palabra a la que se incorpora, sin guion ni espacio intermedios”.

Consequently, the examples found in the corpus show a foreignization of the Spanish orthographical patterns of prefixed items and neoclassical compounds. However, we cannot really claim them as being Anglicisms, although the number of examples retrieved from the corpus is quite significant. One of the possible reasons for imitating the pattern is to create a parallelism with the English language in the consumer. This linguistic creativity is used to resemble English patterns and can have a visual effect provoking an attention-grabbing effect in consumers.

Besides, other examples included in the list are word creations resembling English morphological patterns, like the pattern ‘prefix+hyphen+root’, for example: the English equivalent of *anti-radical* is *free radical*, but the former has been coined in Spanish; and the English expression for *anti-bolsas* is ‘eye cream for bags’, but the derivative has been created in Spanish. In some other cases, *antienvjecimiento* and *antiedad* are recorded in Rodríguez González (2017: 24) without the hyphen too and explaining that both are calques in Spanish.

	Germaine de Capuccini	Natura Bissé	SkinClinic	Farma Dorsch
Anti-	Anti-arrugas	Anti-arrugas		Anti-arrugas
				Anti-bolsas
	Anti-edad /antiedad	Anti-edad /antiedad		Anti-edad /antiedad
		Anti-envejecimiento		
		Anti-estrés /antiestrés		
		Anti-estrías		
	Anti-fatiga	Anti-fatiga		Anti-fatiga
		Anti-flacidez		Anti-flacidez
				Anti-irritante
				Anti-ojeras
		Anti-oxidante		Anti-oxidante
	Anti-polución / antipolución			
		Anti-poros		
		Anti-radicales		
				Auto-protegerse
	Bio-	Bio-degradable		
				Bio-estimulación
Bio-tecnológico				
Dermo-		Dermo-compatibles		
		Dermo-cosmética		
Eco-				Eco-sostenible
Extra-	Extra-confort			
	Extra-rica			
Micro-	Micro-capas			
		Micro-circulación	Micro-circulación	
	Micro-colágeno			
	Micro-encapsulado			
	Micro-esferas			
	Micro-exfoliante	Micro-exfoliante		
		Micro-inflamación		
	Micro-inyecciones	Micro-inyecciones		

		Micro-lifting / microlifting		
		Micro- reflectores		
Multi-		Multi-activa		
	Multi-corrección			
	Multi-estratégica			
	Multi-lamelar			
	Multi-protección			
	Multi-reparador			
Post-		Post-baño		
			Post-peelings	
		Post-partos		
	Post-quirúrgicas			
Pre-			Pre-estrés	
Re-	Re-esculpe			
	Re-generación			
Ultra-	Ultra-calmante			
	Ultra-corrector			
	Ultra-dosificado			
	Ultra-eficaz			
	Ultra-fresco			
			Ultra- hidratante	Ultra- hidratante
		Ultra-sensorial		
	Ultra-suave			

Table 3: Examples from the corpus: prefixed items and neoclassical compounds

The processes of derivation from loanwords are also interesting in that they reveal the richness of lexical coinage. The Anglicism *detox* has been completely assimilated in the Spanish language, as users have coined derivatives, such as a verb, a noun and two adjectives: *detoxificar*, *detoxificación*, *detoxificante*, *detoxificador*. We can compare this case with others such as, for example, *turístico* and *liderar*, whereby an English borrowing, *turismo* and *líder*, is combined with a Spanish suffix to form a derivative.

We have found two more items that should be explained: the terms *voluminizar* and *voluminizador*, which are not included in DRAE. We searched the web to check the number of occurrences and the sources in which they are used, and can conclude that both are lexical coinages from the English terms: *volumize* (“Of a product or styling technique: to enhance the thickness of or give body to (hair or eyelashes)” OED) and *volumizing* (“That gives hair (or eyelashes) body or thickness” OED). Nevertheless, the term *volumen* does exist in Spanish with the meaning: “1. Corpulencia o bulto de algo. 2. Magnitud física que expresa la extensión de un cuerpo en tres dimensiones, largo, ancho y alto, y cuya unidad en el sistema internacional es el metro cúbico” (DRAE). However, it seems quite probable that the derivatives have been coined taking the English roots as bases and adding the Spanish morphemes.

Some of the borrowings analyzed in the study already have and others will be transferred to the general lexicon of the language, as these products are commercialized, advertized and popularized, hence known, bought and used by consumers. Because, as

we have already explained, the borrowing process in this specialised field is closely related to the popularity of cosmetics and their impact on the consumers.

5. Conclusion

The impact of English upon Spanish in the field of cosmetology is undeniable and comparable to the Anglicization process in several languages that some authors have already pointed out (Kachru, 1994; Furiassi, Pulcini and Rodríguez, 2012). The progressive Anglicization of Spanish lexis can be corroborated with all the research carried out by many scholars in the last decades, as all the studies mentioned have shown. Moreover, a number of factors explaining the adoption and use of Anglicisms have been explained.

The English language is generally used as a strategy to provide Spanish speakers with a certain taste of modernity and other positive values as it is seen as an attractive and fashionable language. Thus, attention-getting can be regarded as one of the most important techniques for using Anglicisms and pseudo-Anglicisms in the information included in the leaflets of cosmetic products, which also serves as a basis for the advertisements prepared to sell those products. Besides, borrowings and the influence of English orthographical word-formation patterns on Spanish words found in cosmetics information may help to reach the goals of sales promotion and may have the intention of persuading consumers to buy these products. Nevertheless, it would be necessary to carry out a sociological study to examine if the influence of the English language on the Spanish language has an impact on consumers.

The tendency to use technical or pseudo-scientific terms that originate from English, sometimes using prefixes or Latin and Greek combining forms and creating lexical items that do not exist in English or resembling English word-formation orthographical rules (i.e. *anti-radical*, *anti-bolsas*, *bio-degradable*, *eco-sostenible*, *dermo-cosmética*) has been observed in our study and in previous ones (Rodríguez Medina, 2016a: 159-160). The use of English fulfills mainly socio-psychological effects, and is commonly employed to evoke connotations of professionalism and medical expertise. In fact, it is undeniable that today society associates the English language to modernity and fashion or technical and economic progress, and “the prestige associated with English, sheer snobbery, linguistic interference or even stylistic creativity” (Oncins Martínez, 2012: 235). Hence, the phenomenon of borrowing should be addressed from a linguistic and psychological perspective in order to cover the angles of the process.

Our study indicates that a considerable number of Anglicisms and false Anglicisms are being used in the cosmetic field. When comparing our results (35,29% false Anglicisms) with a previous study (18,6% false Anglicisms) (Rodríguez Medina, 2016a: 162), the examples of false Anglicisms have significantly increased (16,69%). In fact, we do agree with the idea that “false anglicisms are an endless source of information on morphological and semantic evolutions in language contact and on sociolinguistic factors for language borrowing” (Balteiro and Campos, 2012: 250). Even so, we consider that if the ingredient lists and the name of the products had been part of the corpus in this study,

the results would have been different and an attempting hypothesis could be that an exponential increase in the number of false Anglicisms would have been expected. Moreover, the number of Anglicisms is lower than in the previous mentioned study (Rodríguez Medina, 2016a: 162); therefore, extensive research and a larger corpus will be necessary to examine “the prestige of English in today’s world has contributed to the fact that most Anglicisms are used in their original form” (Rodríguez González, 2012: 295).

Future research is needed as in this work we have not examined the degree of acceptance (Görlach, 2001: xxiv) and the gender assignment of the Anglicisms adopted in Spanish. Additionally, part of the information from the leaflets was not used. Therefore, we consider it would be worth analyzing the names of the products and carrying out a longitudinal study examining whether the names have changed due to the influence of the English language, for example, looking at the cosmetic products advertisements in Spanish magazines from the second half of the 20th century.

Furthermore, we would like to analyze all the cosmetic products, not only the facial ones, produced and distributed by the four cosmetic brands in order to offer a complete analysis of the increasing influence of English on the field. This study could offer us more data to be able to show progressive growth in the number of Anglicisms and false Anglicisms, that is, the Anglicization process of Spanish lexis.

Notes

1. <https://www.cosmeticseurope.eu/cosmetic-products/understanding-label/>
2. http://www.dail-software.com/help/9_en/index.html
3. See appendix B.
4. <http://www3.uah.es/anglicor/index.php>
5. We have consulted the online versions of DRAE, DPD, OED, OLD and LDCE; and the paper versions of NDAng, DEAng and GDAng.
6. These two dictionaries have been used for our research because they are prescriptive and establish the standard use of the terms consulted.
7. This dictionary has been added for this study, as is not included in our database (lexicographical information).
8. It is important to mention that we have not included in the study of the field of Cosmetology the lists of ingredients included in the leaflet and the names of the products, which may have changed partially the results.
9. <http://www.lafiebredelamoda.com/belleza-cuidado-personal/complementos-de-belleza-tratamientos-efecto-flash>
10. <https://knowledge.ulprospector.com/692/pcc-soft-focus-technology/>
11. <https://bhskin.com/stress-and-your-skin/>
12. Even though, Real Academia de la Lengua considers the term *relax* an Anglicism from the English verb *relax*, another possible interpretation is to classify it as a false Anglicism deriving from the English noun *relaxation*.
13. Real Academia de la Lengua explains that the final etymon in the borrowing process is the English language. This type has been recorded with both spellings in our corpus.

14. <http://blog.nutritienda.com/trucos-de-belleza-conoces-el-efecto-flash/>
15. Even one of the senses of the verb *lift* is recorded in OED as being synonym of *facelift*.
16. <https://www.healthgrades.com/symptoms/peeling-skin>
17. <http://www.ost-cosmetics.com/about-skin-care/anti-aging-serum.html>
18. “Stuffy; unventilated, not open to the air. Not using or breathing air. Without air; lacking an atmosphere”.
19. “Stuffy; not ventilated. Without wind or breeze; still”.
20. Airless paint sprayers: <http://www.portlandcompressor.com/airless/index.aspx>
21. <http://airlessbottles.com/>
22. <http://www.fundeu.es/>

References

- Alfaro, R. (1970): *Diccionario de anglicismos*. Madrid: Gredos.
- Balteiro Fernández, Isabel (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): 23-52.
- Balteiro Fernández, Isabel and Miguel Ángel Campos Pardillos (2012): “False anglicisms in the Spanish language of fashion and beauty”. *Ibérica*, 24: 233-260.
- Bhatia, Tej (2001): “Language mixing in global advertising”. In E. Thumboo, ed., *The Three Circles of English*. Singapore: University Press, 195-215.
- Bhatia, Tej (2006): “World Englishes in global advertising”. In B. B. Kachru, Y. Kachru and C. Nelson, eds., *The Handbook of World Englishes*. Oxford: Blackwell Publishing, 601-619.
- De la Cruz Cabanillas, Isabel and Cristina Tejedor Martínez (2012): “Email or *correo electrónico*? Anglicisms in Spain”. *RESLA*, 25: 95-117.
- De la Cruz Cabanillas, Isabel, Tejedor Martínez, Cristina, Díez Prados, Mercedes y Esperanza Cerdá Redondo (2005): “English loanwords in Spanish computer language”. *English for Specific Purposes*, 26(1): 52-78.
- Fundeu-BBVA (2017): *Fundación del Español Urgente*. Available at: <http://www.fundeu.es/>
- Furiassi, Cristiano (2010): *False Anglicisms in Italian*. Milan: Polimetrica.
- Furiassi, Cristiano, Pulcini, Virginia and Félix Rodríguez González, eds. (2012): *The Anglicization of European Lexis*. Amsterdam: John Benjamins.
- Gimeno Menéndez, Francisco and María Victoria Gimeno Menéndez (2003): *El desplazamiento lingüístico del español por el inglés*. Madrid: Ed. Cátedra.
- Görlach, Manfred, ed. (2001): *A Dictionary of European Anglicisms. A Usage Dictionary of Anglicisms in Sixteen European Languages*. Oxford: Oxford University Press. [DEAng]
- Görlach, Manfred (2002): *English in Europe*. Oxford: Oxford University Press.
- Gómez Capuz, Juan (1998): *El préstamo lingüístico. Conceptos, problemas y métodos*. Anejo XIX de la Revista *Cuadernos de Filología*. Valencia: Univ. Valencia.
- Gómez Capuz, Juan (2000): *Anglicismos léxicos en el español coloquial*. Cádiz: Universidad de Cádiz.
- Gómez Capuz, Juan (2004): *Préstamos del español: lengua y sociedad*. Madrid: Arco Libros.
- Gottlieb, Henrik and Cristiano Furiassi (2015): “Getting the grips with false loans and pseudo-Anglicisms”. In C. Furiassi and H. Gottlieb, eds., *Pseudo-English: Studies on False Anglicisms in Europe*. Boston/Berlin: De Gruyter Mouton, 3-33.
- Haugen, Einar (1950): “The analysis of linguistic borrowing”. *Language*, 26: 210-231.

- Hsu, Jia-Ling (2008): "Globalization and English mixing in advertising in Taiwan: Its discourse domains, linguistic patterns, cultural constraints, localized creativity, and socio-psychological effects". *Journal of Creative Communications*, 3: 155-183.
- Kachru, Braj B. (1994): "Englishization and contact linguistics". *World Englishes*, 13(2): 135-154.
- King, Henry (2013): *The Benefits of Cosmetics/Toiletries. A Cosmetic Chemist's View*. Australian Society of Cosmetic Chemists. Available at: <https://ascc.com.au/the-benefits-of-cosmeticstoiletries/>
- Longman Dictionary of Contemporary English* (2014): Pearson Longman. Available at: <https://www.ldoceonline.com/> [LDCE]
- Lorenzo Criado, Emilio (1994): *El español lengua en ebullición*. Madrid: Gredos.
- Lorenzo Criado, Emilio (1995): Anglicisms. In Manuel Seco and Gregorio Salvador (eds.). *La lengua española hoy*. Madrid: Fundación Juan March, 165-174.
- Lorenzo Criado, Emilio (1996): *Anglicismos Hispánicos*. Madrid: Gredos.
- Martin, Elizabeth (2002): "Mixing English in French advertising". *World Englishes*, 21: 375-402.
- Merriam Webster Dictionary* (2016): Merriam Webster. Available at: <https://www.merriam-webster.com/>
- Oncins Martínez, José Luis (2009): "Towards a corpus-based analysis of anglicisms in Spanish: A case study". *International Journal of English Studies*, 9(3): 115-132.
- Oncins Martínez, José Luis (2012): "Newly-coined Anglicisms in contemporary Spanish. A corpus-based approach". In C. Furiassi, V. Pulcini and F. Rodríguez, eds., *The Anglicization of European Lexis*. Amsterdam/Philadelphia: John Benjamins Publishing Company, 217-238.
- Oxford University Press (2000): *Oxford English Dictionary*. Available at: <http://www.oed.com/> [OED]
- Oxford University Press (2017): *Oxford Living Dictionary*. Available at: <https://en.oxforddictionaries.com/> [OLD]
- Pratt, Chris (1980): *El anglicismo en el español peninsular contemporáneo*. Madrid: Gredos.
- Pulcini, Virginia, Cristiano Furiassi, and Félix Rodríguez González (2012): "The lexical influence of English in European languages: From words to phraseology". In C. Furiassi, V. Pulcini and F. Rodríguez, eds., *The Anglicization of European Lexis*. Amsterdam/Philadelphia: John Benjamins Publishing Company, 1-24.
- Real Academia Española (2001): *Diccionario de la Lengua Española*. Madrid: Espasa Calpe. Available at: www.rae.es [RAE]
- Real Academia Española (2005): *Diccionario Panhispánico de dudas*. Available at: <http://www.rae.es/recursos/diccionarios/dpd> [DPD]
- Real Academia de la Lengua. (2010): *Ortografía de la lengua española*. Available at: <http://aplica.rae.es/grweb/cgi-bin/buscar.cgi>
- Rodríguez González, Félix (1996): "Functions of Anglicisms in contemporary Spanish". *Cahiers de Lexicologie*, 68(1): 107-128.
- Rodríguez González, Félix (1999): "Anglicisms in contemporary Spanish". *Atlantis*, 21: 103-139.
- Rodríguez González, Félix (2002a): "Spanish". In M. Görlach, ed., *English in Europe*. Oxford: Oxford University Press, 128-150.
- Rodríguez González, Félix (2002b): "Anglicismos y calcos en el español actual". In Felix San Vicente, ed., *L'inglese e le altre lingue europee – Studi sull'interferenza linguistica*. Bologna: CLUEB, 149-169.

- Rodríguez González, Félix (2003): “Orientaciones en torno a la elaboración de un corpus de anglicismos”. In M.T. Echenique and J. Sánchez, eds., *Lexicografía y lexicología en Europa y América. Homenaje a Günther Haensch en su 80 aniversario*. Gredos: Madrid, 559-575.
- Rodríguez González, Félix (2012): “Anglicismos en el mundo del deporte: variación lingüística y sociolingüística”. *Boletín de La Real Academia Española (BRAE)*, Tomo XCII (Cuaderno CCCVI), 317-341.
- Rodríguez González, Félix (2017): *Gran diccionario de anglicismos*. Madrid: Arco/Libros-La Muralla. [NAng]
- Rodríguez González, Félix and Antonio Lillo Buades (1997): *Nuevo Diccionario de Anglicismos*. Madrid: Editorial Gredos. [GDAng]
- Rodríguez Medina, María Jesús (2016a): “The use of anglicisms in Spanish television commercials of cosmetics, hygiene and personal care products”. *Hermes. Journal of Language and Communication in Business*, 55: 157-169.
- Rodríguez Medina, María Jesús (2016b): “An approach to the study of the use of English in the activities of Spanish gyms”. *Spanish in Context*, 13(1): 128-148.
- Rodríguez Segura, Delia (1999): *Panorama del anglicismo en español. Presencia y uso en los medios*. Almería: Universidad de Almería.
- The Watchtower (2012): *The Use of Cosmetics in Bible Times*. Available in: <https://www.jw.org/en/publications/magazines/wp20121201/use-of-cosmetics-in-the-bible/>
- Weinrich, Uriel (1953): *Languages in Contact. Findings and Problems*. Berlin: Mouton de Gruyter.

Appendices

A. Lexical items included in the four cosmetic brands

	Germaine de Capuccini	Natura Bissé	SkinClinic	Farma Dorsch
acti-aging				X 1
after-shave		X 1	X 3	
after-sun		X 1	X 1	
airless				X 10
aminoessence Cocktail		X 7		
anti-aging			X 2	
antiestrés		X 1		
booster	X 3			
coach	X 2			
cocktail		X 2		
comfort	X 1			
confort	X 37	X 14	X 5	
detox	X 4	X 1		
detox-diet complex		X 2		
detoxificar	X 1			
detoxificación	X 1			
detoxificante	X 4	X 1		
detoxificador				X 2
estrés	X 11	X 9		
eyeliner		X 1		

flash	X 2	X 1	X 6	X 4
film	X 1		X 2	
ginger				X 2
hydractivas	X 10			
hydracure	X 4			
hydra-flow		X 3		
inflamm-aging		X 1		
kit	X 1	X 1	X 4	
láser/laser	X 1	X 2	X 7	X 3
lifting	X 11	X 18	X 15	X 6
lipo-filling	X 5			
no-stress	X 2			
oil-free	X 1		X 24	
packaging	X 1			
parabenos/parabenos	X 1	X 12		X 24
peeling	X 1	X 2	X 20	X 5
phyto-complejo	X 5			
relax		X 1		
resiliencia cutánea	X 3			
roll-on	X 2			
scrub			X 4	
sérum	X 14	X 29	X 11	Suero/suerum 5
Stress-reduction	X 2			
tecnología 3D		X 2		
tecnología soft focus	X 1			
test	X 1	X 2	X 1	X 1
testado	X 12	X 4		X 1
voluminizador	X 2			
voluminizar				X 1
waterproof	X 1			
water-oil-free			X 1	
TOTAL lexical types: 50				
	148	122	106	65
TOTAL tokens:	440			

B. Anglicisms: contexts and information from sources

	Contexts from corpus	Information from Spanish dictionaries	Information from English dictionaries and web pages
<i>acti-aging</i>	TECNOLOGÍA CORRECTIVA ACTI-AGING CON ALFAHIDROXIACIDOS. (FD)		
<i>after-shave/ aftershave</i>	Para pieles atópicas y como aftershave. (SC) Es el after-shave ideal para las pieles masculinas tras el afeitado diario. (NB)	En inglés, <i>after-shave lotion</i> (NDAng)	Any of various preparations and products applied to the skin after shaving; (now esp.) an astringent scented lotion applied in this way.(OED)
<i>after-sun/ aftersun</i>	regenerador celular, cicatrizante, recuperador y aftersun. (SC) Es un after-sun ideal para tu piel. (NB)	En inglés, <i>after-sun lotion</i> (NDAng)	A product applied to the skin after exposure to the sun to help reduce sunburn, rehydrate the skin, or enhance tanning. (OED)
<i>anti-estrés</i>	el cual proporciona un excepcional efecto calmante y antiestrés. (NB)		Stress-free (OD) pertaining to or possessing freedom from mechanical or biological stress.(OED)

<i>airless</i>	Innovador envase con sistema al vacío (airless). (FD) Envase con sistema “airless”. Una vez levantada la tapa encontrará el dosificador airless. (FD)		Stuffy; unventilated, not open to the air. Not using or breathing air. Without air; lacking an atmosphere. (OED) Stuffy; not ventilated. Without wind or breeze; still. (OLD)
<i>aminoessence Cocktail</i>	El exclusivo Aminoessence Cocktail es un completísimo cóctel de aminoácidos que constituye un verdadero tratamiento de choque para la firmeza y redensificación de la piel. (NB)		Amino Essence contains all of the essential amino acids to rapidly stimulate the build up of muscle mass, aswell as supporting the repair and recovery of muscles after intense training. http://www.fatbirds.co.uk/298597/products/powerbar-amino-essence-drink-400g-lemon.aspx Tea & Amino Essence Moisturizing Facial Mask
<i>anti-aging</i>	Sérum concentrado anti-aging y luminosidad. Recupera la piel opaca en segundos. (SC)	Antienvejecimiento es el calco más literal, pero en tiempos recientes se prefiere antiedad, por resultar más eufemístico. (GDAnG)	Not included in OED. Adj. (of a product or technique) designed to prevent the appearance of getting older. (Oxford dictionaries)
<i>booster</i>	<i>Booster Generador de Luz y Juventud.</i> (GC)	≠ different meaning. Propulsor utilizado en lanzadores y aviones militares para proporcionar un mayor empuje en el despegue. Del inglés, lit. ‘propulsor’.	Med. A dose or injection of a substance that increases or prolongs the effectiveness of an earlier dose or injection. (OED)
<i>coach</i>	Un auténtico “coach” para la piel. (GC)	Persona encargada de prestar apoyo a otra para alcanzar determinados objetivos (físicos, pero también psicológicos, laborales, etc.), como si se tratara de un entrenador. (GDAnG)	To prepare (a candidate) for an examination; to instruct in special subjects; to tutor; also, to train for an athletic contest, as a boat-race. (OED)
<i>cocktail</i>	Natura Bissé ha dado con el cocktail perfecto para que sientas la vida con intensidad. (NB)	La voz inglesa <i>cocktail</i> se ha adaptado al español con dos acentuaciones, ambas válidas. La forma llana <i>cóctel</i> (pl. <i>cócteles</i>), que refleja la pronunciación etimológica, es la única usada en España. (DPD)	Any combination of ingredients, factors, or circumstances. (OED)
<i>comfort</i>	otra más ligera, Comfort, con las mismas propiedades para pieles normales. (GC)		State of physical and material well-being, with freedom from pain and trouble, and satisfaction of bodily needs; the condition of being comfortable. (OED)
<i>confort</i>	Su textura elegante y suntuosa proporciona una sensación inmediata de confort y nutrición a la piel. (NB)	Confort. Del fr. confort, y este del ingl. comfort. l. m. Bienestar o comodidad material. (DRAE)	

<i>detox</i>	Nueva Fórmula detox iluminadora – bolsas y ojas que combate el “cansancio celular”. (GC) Hidratante detox con color. (NB)	El anglicismo adaptado <i>detox</i> y su traducción desintoxicante son ambas válidas para referirse a este tipo de dieta o menú. El término inglés <i>detox</i> se ha asentando en el campo de la nutrición y la dietética (Fundeu) Referido a un proceso de desintoxicación o purificación, especialmente a la eliminación de toxinas mediante zumos. El método se ha popularizado hoy entre los que cuidan de su salud por influencia de nutricionistas y dietistas. (GDAnG)	= detoxification n. at <i>detoxicate</i> v. Derivatives Also: a detoxification centre. Frequently attrib., esp. of a place used for the treatment of alcoholics or drug addicts. Detoxicate. To deprive of poisonous qualities. (OED) 2 when you do not eat solid food or only drink special liquids for a period of time, which is thought to remove harmful substances from your body. (LDCE)
<i>detox-diet complex</i>	Incorpora un complejo purificante, Detox-Diet Complex con extractos de berro, té blanco y té verde, que desintoxican la piel y la protegen frente a los radicales libres. (NB)		
<i>detoxificar</i>	Revitaliza, detoxifica e ilumina. (GC)		
<i>detoxificación</i>	supone un notable avance en la detoxificación celular. (GC)		
<i>detoxificante</i>	Su poder antiarrugas y detoxificante, reafirma la piel de los párpados. (NB)		
<i>detoxificador</i>	Antioxidante, detoxificador y reestructurador cutáneo. (FD)		
<i>estrés</i>	Para todo tipo de pieles, especialmente fatigadas y expuestas a un estrés externo extremo. (GC) la refuerza frente al estrés diario y previene la aparición prematura de signos de la edad. (NB)	Del ingl. <i>stress</i> . l. m. Tensión provocada por situaciones agobiantes que originan reacciones psicosomáticas o trastornos psicológicos a veces graves. (DRAE)	An adverse circumstance that disturbs, or is likely to disturb, the normal physiological or psychological functioning of an individual; such circumstances collectively. Also, the disturbed state that results. (OED)
<i>eyeliner</i>	Con un bastoncillo de algodón humedecido en este producto puedes retocar tu eyeliner o tu sombra de ojos. (NB)	Producto normalmente líquido que se emplea, con la ayuda de un pincel, para delinear el borde de los párpados. (En inglés hay tres clases: <i>pencil eyeliner</i> – el más corriente y fácil-, <i>liquid eyeliner</i> y <i>gel eyeliner</i>). (GDAnG)	A cosmetic applied in a line bordering the eye, generally in order to accentuate it; (also) a brush or pencil for applying this. (OED)
<i>flash</i>	Aplicar una pequeña cantidad para un efecto flash. (SC) Hidrata y revitaliza la piel con un efecto flash. (FD)	Impresión repentina y placentera. La connotación del placer se basa en su asociación con la acepción 7, y su instantaneidad se refuerza al evocar también la visión de una imagen como la	Burst of light or flame. (OED)

		producida por el flash de la fotografía. (GDAng)	
<i>film</i>	Despegar el film de plástico y aplicar sobre la piel impregnada con sustancias cosméticas. (SC)	Película muy fina de plástico u otro material utilizada para envolver o embalar algunos productos. (GDAng)	Material in the form of a very thin flexible sheet. (OLD)
<i>ginger</i>	Extracto de Ginger o Jengibre orgánico. (FD)		
<i>hydractivas</i>	Una selección de activos sumamente cuidada convierte su fórmula en hydractiva, 100% actividad hidratante. (GC)	Hidro- 1. elem. compos. Significa 'agua'. (DRAE)	Miscellaneous terms, in which hydro- has the sense of 'water'. (OED) Hydra- Of or belonging to a hydra, hydra-like; having as many heads, or as difficult to extirpate, as the Lernaean hydra. (OED)
<i>hydracure</i>	Hydracure realiza una máxima actividad hidratante. (GC)		Cure: A substance or treatment that cures a disease or condition. (OLD)
<i>hydra-flow</i>	el complejo Hydra-Flow que aporta hidratación y frescor de forma inmediata. (NB)		Hydrate: A compound of water with another compound or an element. (OED)
<i>inflamm-aging</i>	las consecuencias del inflamm-aging: una micro-inflamación silenciosa. (NB)	Unión de los vocablos ingleses "inflammation" (inflamación) y "ageing" (envejecimiento). Este fenómeno describe cómo los procesos inflamatorios conducen a un envejecimiento prematuro de la piel. http://www.montibello.com/esp/in-beauty/2640-inflamm-aging-combate-el-envejecimiento-prematuro-de-las-pieles-sensibles	The primary feature of inflamm-aging is an increase in the body's proinflammatory status with advancing age https://www.hindawi.com/journals/jir/2016/8426874/ As the concept of inflammaging goes mainstream, it may result in a widespread change in the professional skin care industry, transforming everything from the ingredients that professionals use to the methods by which they use them. http://www.skininc.com/skinscience/physiology/Inflammaging-Changing-the-Face-of-Skin-Care-185422772.html
<i>kit</i>	Skinclinic ha ideado este kit de tratamiento facial para cuidar y prevenir la piel del envejecimiento y protegerla del sol diario. (SC)	Del ingl. <i>kit</i> , y este del neerl. <i>kit</i> . 1. m. Conjunto de productos y utensilios suficientes para conseguir un determinado fin, que se comercializan como una unidad. (DRAE)	A set of articles or equipment needed for a specific purpose. (OLD)
<i>láser/laser</i>	Alivio de cualquier sensación de la piel después de tratamientos láser, depilación. (SC) Calma de forma inmediata la sensación de irritación en la piel ocasionada por el sol, quemaduras leves, afeitado, depilación, sequedad, tratamientos con láser. (FD)	Del ingl. <i>laser</i> , acrón. de <i>light amplification by stimulated emission of radiation</i> 'amplificación de luz mediante emisión inducida de radiación'. 1. m. Dispositivo electrónico que, basado en una emisión inducida, amplifica de manera extraordinaria un haz de luz monocromático y coherente. 2. m. Haz de luz emitido por un láser. (DRAE)	Any device that is capable of emitting a very intense, narrow, parallel beam of highly monochromatic and coherent light (or other electromagnetic radiation), either continuously or in pulses, and operates by using light to stimulate the emission of more light of the same wavelength and phase by atoms or molecules that have been excited by some means. (OED)
<i>lifting</i>	Acción lifting mirada despierta. (GC) Conseguirás, de inmediato una piel renovada, extraordinariamente luminosa y con un espectacular efecto lifting. (NB) que induce un efecto lifting, aumentando el tono muscular y	Voz inglesa que se usa con frecuencia en español con el sentido de 'operación de cirugía estética consistente en el estiramiento de la piel, generalmente de la cara y el cuello, para suprimir	

	favoreciendo la capacidad tensora. (SC) Reafirma la piel produciendo un efecto "lifting" inmediato. (FD)	las arrugas'. Se recomienda sustituirlo por el equivalente español <i>estiramiento (facial)</i> .(DPD)	
<i>lipo-filling</i>	más conocido como el 'hongo de la larga vida', de él se extrae un activo que, gracias a sus propiedades "lipo-filling". (GC)		Lipo-: combining form of Greek λίπος fat, used in various pathological terms, chiefly modern Latin, in Biochem. and other fields.(OED)
<i>no-stress</i>	Aplicar Crema hidratante no-stress sobre la piel limpia y renovada de rostro, cuello y escote. (GC)		
<i>oil-free</i>	Cuidado facial oil-free de la piel grasa o acnéica. (SC)		An oil-free liquid, skin treatment etc contains no oil. (LDCE)
<i>packaging</i>	Sus fórmulas, packaging, etiquetado y publicidad estén acordes con los valores de protección del medio ambiente. (GC)	Material empleado para empaquetar productos. Del inglés, lit. 'embalaje'. (GDAng)	Materials used to wrap or protect goods. (OLD)
<i>parabenes/ parabenos</i>	ALTA TOLERANCIA SIN PERFUME, ALERGENOS NI PARABENES (FD) que no contienen siliconas, parabenes ni colorantes artificiales. (GC) Es un producto libre de alcohol y parabenos. (NB)	Parabeno. Del ingl. paraben, acrón. de para-'para-' y hydroxybenzoic [acid] '[ácido] hidroxibenzoico'. (DRAE)	Any of a group of compounds which are alkyl esters of para-hydroxybenzoic acid and are used as preservatives in pharmaceutical and cosmetic preparations and in the food industry. (OED)
<i>peeling</i>	Mantenimiento tras tratamiento peeling o láser despigmentante. (SC) PEELING EXFOLIANTE E HIDRATANTE CON CÉLULAS MADRE VEGETALES PARA CARA Y CUERPO. (FD)	1.Exfoliación suave de la superficie epidérmica (...) En inglés se emplea <i>exfoliation</i> . 2.Producto empleado en este tratamiento. En inglés se emplean otros términos como <i>facial scrub</i> . (GDAng)	
<i>phyto-complejo</i>	PhytoComplejo Stress-Reduction. (GC)		Phyto- with the sense 'of, relating to, or resembling (that of) a plant or plants'. (OED)
<i>relax</i>	Es ideal como tratamiento de relax para cuando sientes la piel de tu rostro fatigada. (NB)	Relajamiento (muscular), relajación, distension, Descanso, bienestar. En inglés se emplea <i>relaxation</i> . (GDAng) Del ingl. Relax.1. m. Relajamiento físico o psíquico producido por ejercicios adecuados o por comodidad, bienestar o cualquier otra causa.	Make or become less tense or anxious. (OLD)
<i>resiliencia (cutánea)</i>	Tratamiento de día que favorece la "óptima resiliencia" cutánea fundamental para preservar la belleza y juventud de la piel. (GC)	Del ingl. resilience, y este der. del lat. resiliens, -entis, part. pres. act. de resillire 'saltar hacia atrás, rebotar', 'replegarse'. 1. f. Capacidad de adaptación de un ser vivo frente a un agente perturbador o un estado o situación adversos.2. f. Capacidad de un material, mecanismo o sistema para recuperar su estado inicial cuando ha cesado la perturbación a la que había estado sometido. (DRAE)	Elasticity; the power of resuming an original shape or position after compression, bending, etc.(OED)

<i>roll-on</i>	Su aplicador en roll-on maximiza su acción global. (GC)	A deodorant stick, etc. (DEAng)	Of liquid (esp. a deodorant): that is applied by means of a rolling stopper at the mouth of the container. (OED)
<i>scrub</i>	Higiene y limpieza física por efecto scrub sobre la piel. (SC)		A deep-cleaning, mildly abrasive soap. Frequently in facial scrub.(OED) A semi-abrasive cosmetic lotion applied to the face or body in order to cleanse the skin. (OLD)
<i>sérum</i>	Sofisticado sérum concentrado, altamente rejuvenecedor capaz de mantener la salud de nuestras células en perfecto estado. (NB) Suero Antiedad Hidratante y Redensificante para cara y cuello. (FD)	El término es bien acogido en el campo de la cosmética en detrimento de su equivalente español <i>suero</i> , tan ligado al ámbito de la medicina. (GDAng)	Watery animal fluid, normal or morbid; spec. blood-serum, the greenish yellow liquid which separates from the clot when blood coagulates. (OED)
<i>Stress-reduction</i>	PhytoComplejo Stress-Reduction. (GC)		Mindfulness-Based Stress Reduction [http://www.umassmed.edu/cfm/mindfulness-based-programs/]
<i>(tecnología) 3D</i>	Su avanzada tecnología 3D actúa triplemente en la arruga. (NB)		The quality of being three-dimensional. (OLD)
<i>(tecnología) soft focus</i>	Vanguardista tecnología soft focus: millones de microesferas que contienen diminutos corazones de diamante luminiscentes compensan la falta de resplandor del relieve cutáneo y éste adquiere un aspecto radiante. (GC)		Deliberate slight blurring or lack of definition in a photograph or film. (OLD)
<i>test</i>	Dermatológicamente Probado con Test de Irritación y Sensibilización. (FD) * Test de eficacia en 40 voluntarias de 32 años de edad media durante un mes de aplicación mañana y noche. (GC)	Del ingl. <i>test</i> . 1. m. Prueba destinada a evaluar conocimientos o aptitudes, en la cual hay que elegir la respuesta correcta entre varias opciones previamente fijadas. (DRAE)	
<i>testado</i>	Testado en todo tipo de piel, incluyendo piel sensible. (NB) Testado dermatológicamente sobre pieles sensibles. (GC)		
<i>voluminizador</i>	Aplicar ultra-corrector voluminizador labios y contorno mediante suaves lisajes y movimientos circulares en la zona de labio y su contorno. (GC)		
<i>voluminizar</i>	Voluminiza, rellena y disminuye las arrugas. (FD)		
<i>waterproof</i>	Fórmula bi-fásica altamente eficaz que elimina con suavidad y rápidamente el maquillaje de ojos y labios, incluso si éste es de larga duración o waterproof. (GC)	Referido al producto cosmético que resiste al agua. Del inglés waterproof, lit. 'a prueba de agua'. (GDAng)	Of a thing: impervious to water, impermeable. (OED)
<i>water-oil-free</i>	Gel barrera silicónico. Water-oil-free. (SC)		

Exploring the dynamics of English/Spanish codeswitching in a written corpus

M^a Isabel González Cruz
Universidad de Las Palmas de Gran Canaria
isabel.gonzalezcruz@ulpgc.es

ABSTRACT

This investigation is part of a much larger ongoing research project which approaches a corpus of popular romance fiction novels from a multidisciplinary perspective. The paper focuses on the usage of Spanish words and expressions in the English written discourse of two samples of romances taken from the corpus we are compiling for Research Project FFI2014-53962-P. When analyzing the occurrences of Hispanicisms in the samples, we will specifically address the issues of both their forms and the different socio-pragmatic functions that these cases of language switching seem to play. It is only recently that scholars have studied the patterns of codeswitching in literary writing, but, to the best of the author's knowledge, no previous research has focused on codeswitching in this particular subgenre, which has always been doubly stigmatized for being both popular and feminine.

Keywords: English/Spanish codeswitching, written codeswitching, Hispanicisms, popular romance fiction, Canary Islands

1. Introduction

This paper is part of a much larger ongoing research project¹ which approaches a corpus of romance fiction novels from a multidisciplinary perspective. The present study focuses on one particular aspect of the four levels of analyses established for that investigation, namely, the incorporation of Spanish words and expressions into their English written



discourse. We will examine a total of 16 romances taken from the corpus that is being compiled for Project FFI2014-53962-P in order to specifically explore the dynamics of English/Spanish codeswitching, focusing on both their types and the socio-pragmatic or discourse functions that these cases of language mixing² seem to play.

Mixed-language written texts have frequently been found in postcolonial writing and travel literature (Curell and de Uriarte, 2001; González-Cruz and González de la Rosa 2006, 2007; Curell et al. 2009; Anchimbe and Janne, 2011; González-Cruz, 2013). It is only relatively recently that scholars have shown an interest in studying the patterns of codeswitching in other types of literary writing. In fact, in the last decades a number of authors have investigated this phenomenon in a variety of genres, such as narrative (Cortés-Conde and Boxer, 2002), plays and drama (Pfaff and Chávez, 1986; Anderson, 2004; Jonsson, 2010), poetry (Valdés-Fallis, 1976; 1977), short stories (Montes-Granados, 2012) or even stand comedy (Woolard, 1987) and TV series (Giménez-Eguibar and de los Heros, forthcoming). Other works have focused on areas such as the cyberspace (Montes Alcalá, 2005b; Söderqvist, 2009) and popular music (Flores Ohlson, 2008), while authors like Montes-Alcalá (2016) have carried out comparative research about the use of codeswitching in different genres. To the best of our knowledge, no previous research has focused on codeswitching in popular romance fiction novels, a sub-literary form that has always been doubly stigmatized for being both popular and female-oriented (Sánchez-Palencia, 1997). As indicated in the title, our aim here is to explore the dynamics of the use of Hispanicisms in this type of English fictional texts.

The paper is structured as follows. After the introductory section, we will approach the theoretical framework from which we draw on, namely the sociolinguistic study of language codeswitching. With an overview of this field of research we will shape up the socio-pragmatic analysis of the use of Hispanicisms in the selected samples as cases of codeswitching. The novels are briefly depicted in section 3, where the hypothesis and the methodological procedure are also outlined. Section 4 describes the data obtained after analyzing the occurrences of Spanish words and expressions in the texts. First we will focus on their formal types and then we will discuss the kind of functions these Hispanicisms seem to perform. Some concluding remarks will be offered before listing the bibliographical references, which also include the primary sources used.

2. Theoretical framework. On codeswitching

Language switching is one of the most salient manifestations of bilingualism and language contact situations, which involve the alternative use of two (or more) languages. The term *codeswitching*, also spelled *code-switching* (henceforth, CS), was used by Haugen³ (1956) in order to refer to situations in which “a bilingual introduces a completely unassimilated word from another language into his speech” (Haugen, 1956: 40). Defined simply as the alternating use of two languages or language varieties in the same stretch of discourse, at the word, clause or sentence level, CS is viewed as “a natural phenomenon commonly attested in speakers of bilingual communities in which two or more languages are in contact” (Montes-Alcalá, 2001: 715). Actually, CS is fostered by

bilingualism, which is a prerequisite to CS, since it is when bilingual speakers converse that they tend to “integrate linguistic material from both of their languages within the same discourse segment” (Bonvillain, 2008: 320). Far from being a random mixing of two languages indicating a speaker’s internal mental confusion, as it used to be thought of decades ago, CS is now taken to reflect “a skillful manipulation of two language systems for various communicative functions” (Bullock and Toribio, 2009: 4). In fact, it has become an increasingly important field of research and has been the concern of disciplines such as sociolinguistics and psycholinguistics, which represent the two main perspectives from which the issue can be approached, the societal and the individual level. While both angles have contributed to the analysis of CS, they also involve problems, as Gardner-Chloros (1991) explained, due to its multidimensional character. Part of the complexity of the study of language switching emerges from the different connotations and social meanings it carries as well as from the fuzzy edges of the phenomenon, such as the borderline between CS and borrowing (Lipski, 2005; Myers-Scotton, 1992).

Two main types of CS are generally distinguished in the literature: *intersentential CS*, which involves “switches from one language to the other between sentences”, and *intrasentential CS*, which occurs “within the same sentence, from single-morpheme to clause level” (Myers-Scotton, 1992: 4). Appel and Muysken (1987) distinguished a third type of CS which involves an exclamation, a tag or a parenthetical which is expressed in a language other than the language used in the rest of the sentence. Because these *tag-switches* serve as an emblem of the bilingual character of an otherwise monolingual sentence, Poplack (1980) uses the term *emblematic switching* for this type of CS. Although these are the primary switches, Thompson and Lamboy (2012: 53) mention a fourth category at the conversation or discourse level. This fourth type is actually included in the intersentential switches and is influenced by the stylistic features sought after.

Another traditional classification is the one proposed by Lipski (1982: 195), who distinguishes three types of CS: Type I involves the simple insertion of words; while Types II and III respectively coincide with cases of *intersentential* and *intrasentential* CS.⁴ Type II occurs in texts which are “produced in a single language, with switches occurring at phrase/sentence boundaries”. As Lipski (1982: 195) explains, “[i]t is impossible, for such texts, to establish the true bilingual competence of the writer, but the fact that entire propositions are expressed monolingually indicates that the writer sees fit to separate the two linguistic and cultural domains.” In contrast, Type III, which is described by Lipski (1982: 195) as the kind of switch that provides “the richest and most rewarding terrain for literary analysis”, really shows “the degree of integration in the writer’s bilingual grammars”. It involves intrasentential code switches, i.e. the switches occur within the sentence or clause, which requires a much higher level of bilingualism and “presupposes a balanced bilingual grammar” (Montes-Alcalá, 2016: 198) to facilitate a smooth combination of the two systems. This explains why in many studies on literary CS this tends to be the least frequent type (Flores Ohlson, 2008; Montes-Alcalá, 2005a, 2015b; Söderqvist, 2009).

CS has generated “a great deal of pointed discussion in the public domain” (Bullock and Toribio, 2009) but also in the scholarly world, mainly because of its many linguistic, social and cognitive implications, which are in line with “the different implications of switching from a dialect to a standard language and switching registers” (Gardner-Chloros, 1991: 57). As Zentella (1997: 80) put it, “[w]hereas monolinguals adjust by switching phonological, grammatical, and discourse features within one linguistic code, bilinguals alternate between the languages in their linguistic repertoire as well.”

Since CS is felt to be more of a construct, researchers “may reasonably differ as to its definition and scope” (Gardner-Chloros, 1991: 186). It has actually been widely studied from different perspectives by an array of scholars who have shown that bilinguals tend to codeswitch for a variety of social and discursive factors (cf. Poplack, 1980; Lipski 1982, 2005; Appel and Muysken, 1987; Heller, 1988, 1992; Gardner-Chloros, 1991; Eastman, 1992; Myers-Scotton, 1993; Milroy and Muysken, 1995; Tabouret-Keller, 1997; Zentella, 1997; Auer, 1998; Jacobson, 1998, 2001; Poplack and Meechan, 1998; Moyer, 1998; Montes-Alcalá, 2000a, 2000b, 2001, 2005a; 2005b; Callahan, 2004). As Gardner-Chloros (2015) put it, “Code-switching is infinitely more than just throwing in a word from your other language. It can serve to bring emphasis, define the speaker’s identity, be conspiratory or even be used to exercise powers and is usually an attempt to have an effect on the listener”. In addition, CS seems to be governed by “a complicated and as yet not fully delimited set of constraints” (Lipski, 1982: 191). Most studies on CS have underlined its complexity and “its endless psychological and sociological ramifications” (Gardner-Chloros, 1991: 68).

The difficulties to characterize CS definitively are due to the broad range of contact phenomena it comprises, which Bullock and Toribio (2009: 2) summarize as follows:

First, its linguistic manifestations may extend from the insertion of single words to the alternation of languages for larger segments of discourse. Second, it is produced by bilinguals of differing degrees of proficiency who reside in various types of language contact settings, and as a consequence their CS patterns may not be uniform. Finally, it may be deployed for a number of reasons: filling linguistic gaps, expressing ethnic identity, and achieving particular discursive aims, among others. Given these factors, it is not surprising that there exists debate in the literature concerning the precise characterization of CS and how various kinds of language contact varieties are to be classified.

Interestingly, scholars like Giles and Smith (1979) consider CS as part of a linguistic accommodation process which allows speakers to ‘converge’ or ‘diverge’ from their interlocutors. Following this line of thought, Jackson (2014: 142) claims that “language may be used as an identity marker to either draw us closer to or further apart from individuals with a different linguistic and cultural background”, while others, like Tabouret-Keller (1997: 147), conceive it as a way to reinforce the sense of a shared, mixed identity and to maintain the *status quo* of the bilingual situation. This is nothing but evident, as long as CS “can reflect dual identity” (Clyne 2003: 214); in fact, as Lipski (1982: 196) put it, for many bilinguals in the USA, “the dynamic switch to Spanish is fundamentally an identity marker in speakers whose linguistic abilities would allow them

to express themselves equally well in English.” Similarly, Myers-Scotton (1993) discussed the socio-psychological aspect of CS and argued that it can be used to convey intentional meanings of a socio-pragmatic nature. In her view, this socio-psychological significance of switching either languages or linguistic varieties seems to reveal “that the speaker has the multiple identities associated with each of the linguistic varieties involved” (1993: 7). This goes in line with Heller’s (1988: 270-1) assertion that “CS is a strategy which can signal shared culture or be used to create it.” More recently, Pennycook (2012) conceptualizes CS as functioning as an index of affective experiences, while Bonnin (2014) understands it as an alternation between symbolic representations of language.

Several researchers have also described the many important functions served by CS, which will be commented on below. As Gardner-Chloros (1991: 190) explains, it seems obvious that speakers’ motivations for CS are often complex and tend to be made up of several layers which researchers try to disentangle. However, it is also noteworthy that for some bilinguals CS “merely represents another way of speaking” in such a way that often they codeswitch “simply because they can and oftentimes may not be aware that they have done so” (Bullock and Toribio, 2009: 11). Alternatively, in the literature on CS mention is made of cases in which the bilingual speaker may just opt for the language with the shortest or simplest word, the choice being also sometimes due to linguistic ability, vocabulary accessibility or even to occasional memory lapses (Bonvillain, 2008: 321).

On the other hand, it is interesting to note that most of the research dealing with CS has focused on the oral production of bilinguals (Grosjean, 1982; Auer, 1984; Zentella, 1997; Moyer, 1998; Heredia and Altarriba, 2001; Anderson and Toribio, 2007; Lipski, 2014, among others). It is only relatively recently that some scholars have shown an interest in analyzing written CS (Sebba, 2000; Callahan, 2004; Montes-Alcalá, 2005a, 2005b, 2012, 2015; Sebba, Mahootian and Jonsson, 2012; Müller, 2015) or even contrasting the two modes, speech and writing (Lipski, 1982; Montes-Alcalá, 2000a, 2000b, 2001). Thus, Callahan (2004) proves that written CS follows for the most part the same syntactic patterns as its spoken counterpart. She offers an overview of written CS research, and concurs with other scholars (Keller, 1984; Kanellos, 2003) that CS is used by writers to achieve their aims.

Likewise, Sebba, Mahootian and Jonsson (2012) have approached CS in writing as rhetorical practices of contact zones. They illustrate how language mixing in written discourse has occurred historically and how this phenomenon continues in our contemporary time, when interest in this particular subject has increased noticeably. This is also evidenced by the appearance of a special issue of the academic journal *Language and Literature* (2015, vol. 24.3) specifically devoted to CS in literature. This volume has answered the long felt need of more critical work on CS in written discourse which, in contrast to its oral counterpart, had not developed such a considerable amount of literature. It is only in the last decades that scholars have started filling the gap, with the analysis of written CS in a variety of genres. Our aim is to modestly contribute to the

field with this piece of research that explores the dynamics of CS in a sample of popular romance fiction novels set in the Canaries.

3. Corpus description, hypothesis and methodology

As stated above, the romances that will be examined here were taken from the corpus that is being compiled for Research Project FFI2014-53962-P. This interdisciplinary project emerges from previous research on the Anglo-Canarian sociocultural, linguistic and literary relationships, particularly from González-Cruz's (1995, 2002, 2011a, 2011b, 2013, 2014) studies on the wide English bibliography on the Canaries. Popular romance fiction novels were not included in those bibliographical repertoires and have only recently begun to be taken into consideration (González-Cruz, 2015). All the texts in the corpus were published by Harlequin and Mills and Boon female English-writing authors and follow the typical pattern of this denigrated subgenre, i.e., they all tell stories that are centred on love and the couple and conclude with "an emotionally-satisfying and optimistic ending" (cf. *Romance Writers of America*). The novels share one crucial feature: they include a trip and a stay in one or more of the Atlantic islands, which are typically described as a paradise with very different sociocultural traditions. Interestingly, both the narrator and the characters are aware of the sociocultural and linguistic differences and tend to make reference to them. This provides an interesting framework and material to carry out a variety of analyses related to a number of challenging issues, including identity and otherness, gender, paradise discourse, as well as intercultural and even linguistic contact, since the texts tend to be interspersed with Spanish (and Portuguese) terms and expressions. These words and phrases perform different communicative functions and reach a wide international readership who can't but become familiar with them. In turn, this may contribute to their eventual incorporation into the lexical repertoires of the English language, as argued by González-Cruz (2011b; 2017).

For the present study, I selected a total of 16 romances (cf. References) which, just like the rest of the works in the Project corpus, are all set, wholly or partially, in the Canaries. The texts have been divided into two samples on the grounds of their protagonists' origins. Thus, the first one (which we will call *sample A*) includes a total of 8 novels whose protagonists make up a mixed couple, namely a Hispanic (Canarian) hero and an English-speaking heroine who visits or settles on the island temporarily. The second sample (*sample B*) also consists of a total of 8 novels but their protagonists are both of British origin with no family or close social bonds in the Canaries. Since by definition CS and borrowing tend to occur in bilingual situations, such as the ones portrayed in these texts, we expected the romances in *sample A* to include a higher number of Hispanicisms than those in *sample B*, whose protagonists are both British citizens visiting the islands either as mere tourists or for business. With this hypothesis in mind, we carried out our study of the presence and the form and functions of the Hispanicisms in the novels.

As for the methodology, it consisted in carefully reading the selected texts in order to manually log all the occurrences of Hispanicisms, used either by the narrator or by the

characters. As stated, two main perspectives were adopted for the description and analysis of these Hispanicisms, namely their forms and the functions they seem to play in the written discourse in English. For the former, we rely on the distinction between lexical and syntactic CS, depending on whether they respectively involve single words, or phrases and sentences. In this type of analysis Lipski's (1982) well-known typology for CS was employed.

In the analysis of the functions performed by the Hispanicisms in the text, we will also focus on the potential that they seem to have as a means of pragmatic marking of the English discourse of the novels by creating associations and meaning extensions (Khoutyz, 2009), since we are totally aware of "the symbolic significance of language choice" (Clyne, 2003: 42). In this respect, we rely also on Culpeper's (2014) claim that "the words of a text create a particular impression of a character in the reader's mind". This socio-pragmatic viewpoint will be highlighted in the analysis of the cases of CS in our samples.

4. Results and discussion

The results of both the formal and the functional analyses of the Hispanicisms in the two samples will be presented separately here. Thus, in sub-section 4.1 we will study the data obtained from the perspective of their form, following Lipski's (1982) typology. Then, in 4.2 we will focus on the functions that the Hispanicisms collected seem to perform in the English texts. For this purpose, we resorted to the typologies offered by several authors as a guide to design our own classification.

4.1. Lexical vs. syntactic CS: Lipski's (1982) typology

A total of 264 different Hispanicisms were found in the 16 romances studied, including words, phrases and sentences. Neither toponyms nor the many occurrences of the same Hispanicism were considered for the count. Tables 1 and 2 below provide the data obtained, classified by sample. When comparing the figures in both tables, we can notice that there is a tendency for the works in sample A to include a higher number of Hispanicisms than those in sample B, except for two romances in the latter, those by Airlie (1958) and Danbury (1973) in which a considerable number of Spanish words are also employed. Having read all the novels, we are in a position to try to give an explanation for these exceptional cases: in Airlie's (1958) and Danbury's (1973) texts, even though the male protagonists are native English speakers, they have been long established in the Canaries, where they have close social bonds, thus becoming relatively Spanishized or Canarianized. In addition, the plots in these two novels develop mainly in the islands, while in the rest of the romances in this sample, part of the action takes place in other locations, both in Britain and in continental Spain. These factors may explain the contrast between the considerable number of Spanish words inserted in these two novels (and in the ones in sample A) and the relatively few Hispanicisms included in the other works in sample B.

As expected, this seems to indicate that the fact that the protagonists form a mixed couple (invariably a Spanish hero and an English-speaking heroine) implies a greater involvement with the Hispanic context, which in turn determines the occurrence of more Hispanicisms in the texts. This can be interpreted as a literary strategy used by the authors in order to represent the characters' bilingual interactions and to recreate more vividly the bilingual context in which the plots are developed.

As shown in Tables 1 and 2 below, the majority of the Hispanicisms collected were cases of lexical CS. The figures also prove that switches involving phrases were more frequent than those which meant the insertion of sentences.

Authors	Insertion of different Spanish words	Insertion of different Spanish phrases	Insertion of different Spanish sentences
Arbor (1967)	47	15	0
Britt (1977)	24	8	0
Corrie (1980)	20	4	1
Lane (1978)	33	7	4
Mayo (1988)	20	3	1
Mayo (1994)	31	8	5
MacLeod (1982)	48	7	3
Thorpe (1973)	38	10	1

Table 1: Total number⁵ of Spanish words, phrases and sentences found in sample A (mixed couples).

Authors	Insertion of different Spanish words	Insertion of different Spanish phrases	Insertion of different Spanish sentences
Airlie (1958)	38	1	1
Danbury (1973)	29	12	0
Field (2004)	3	0	0
Jameson (1987)	6	0	0
Mayo (1990)	3	0	0
Mayo (1992)	5	2	1
Olfield (1986)	12	1	0
Peake (1983)	3	0	0

Table 2: Total number of Spanish words, phrases and sentences found in sample B (British couples)

Following Lipski's widely-accepted classification for written CS, we must say that many of the romances compiled for the Project belong to Type I only, i.e., they can be described as monolingual texts "with a handful of L2 words thrown in for flavor" (Lipski, 1982: 195). However, in the two selected samples (particularly in sample A) a considerable number of cases of both Type I and Type II have been found, the former

being more numerous than the latter, as suggested in Tables 1 and 2 above. However, no real cases of Type III could be identified.

Below we offer examples⁶ of the two types of CS that were used in the texts. As stated above, Type I, merely involves “the insertion of individual items for a variety of reasons and effects”, as illustrated in the following utterances:

- (1) I still do, *querida*. (Lane, 1978: 185)
- (2) They found rooms in a *pension*, and Rhiannon was too tired... (Mayo, 1988: 201)
- (3) Lucía came out to the *patio* to join them. (MacLeod, 1982: 151)
- (4) Señora Perestello had only just emerged from her own *siesta*. (Thorpe, 1973: 123)
- (5) They had a simple meal of *paella*, cheese, local wine and coffee. (Lane, 1978: 56)
- (6) ‘But you remember his *novia*?’ (Thorpe, 1973: 123)
- (7) ‘She came to return your *mantilla*, which I’d left at the bungalow.’ (MacLeod, 1982: 158)
- (8) ‘I’m very much looking forward to the *fiesta*’. (Thorpe, 1973: 89)
- (9) ... the sound of voices from the *salón* below... (MacLeod, 1982: 156)
- (10) ... full of women chattering over their *merienda*. (MacLeod, 1982: 26)
- (11) ... a choice of *mariscos*, *pulпитos*, *gambas* and many fishes. (Oldfield, 1986: 61)
- (12) There is a *supermercado* only (Jameson, 1987: 73)
- (13) ... I will go and have a word with the *policía*, even though... (Mayo, 1992: 24)
- (14) ... encouraged to be *macho* by indulgent parents. (MacLeod, 1982: 63)

In turn, Type II typically involves switches that occur at the phrase or sentence level, offering “two complete sets of logical propositions”. According to Lipski (1982: 195), it usually “entails a shift of total domain of discourse”, although this does not always seem to be the case in the following examples taken from our samples:

- (15) ‘...but I see your-er-your *hermana*. *Comprende usted?*’ (Mayo, 1992: 118)
- (16) ...her dark eyes aglow. ‘*Te quiero mucho, Phillip!*’ she sighed... (Airlie, 1958: 49)
- (17) ... she heard Manolo’s excited voice. ‘Papá, Papá, donde estás?’ (Mayo, 1994: 152)
- (18) ‘*Yo comprendo!* It is a good thing, is it not, to laugh together and be happy?’ (MacLeod, 1982: 111)
- (19) ‘*Te quiero*’. This time she was the one to murmur those wonderful Spanish words (Lane, 1978: 185)
- (20) ‘*Perdóneme*, I thought I was doing you a favour. If you don’t want to enjoy the *carnaval*, then so be it. *Es tuya la perdida, no mia*’. (Mayo, 1994: 70)
- (21) ‘If you wish, but it will be of no consequence either way. *Buenas noches, Cathy. Lo siento mucho!*’ (MacLeod, 1982: 113)
- (22) ‘Tanya, this is my son, Manolo. Manolo, *esta es la señora de quien te hablé*.’ The boy frowned. ‘*No comprendo, Papá.*’ (Mayo, 1994: 87)

- (23) ‘*Buenas noches everyone! Velocidad moderada, Ramón!*’ (MacLeod, 1982: 134)
 (24) ‘*Tomorrow the fiesta. Adiós, chica!*’ (Thorpe, 1973: 87)

4.2. The sociopragmatic functions of CS

Here we will focus on the variety of linguistic and interactional functions that CS can perform. These are seen as functions of a socio-pragmatic nature, as they have to do with the speaker’s (or writer’s) communicative intention, that is, whatever the speaker/writer wants to express through CS and the social functions CS can play, as well as with how CS affects the hearer’s (or reader’s) interpretation. Before offering our findings, we will make reference to some previous research into this issue, which proved to be helpful for the classification of our data.

4.2.1. Previous studies

There is no single classification or a specific number of functions established for CS. Actually, an array of scholars have listed several stylistic or sociopragmatic functions it can perform, usually to suit the findings of their studies. This is nothing but natural since, as Clyne (2003: 241) put it, following Muysken (2000),

A researcher’s data, experience and interests exercise an influence on the models they develop. With the great expansion of research in this field, language contact studies are served by a multiplicity of models each of which captures part of the picture and contributes to the totality.

Thus, Valdés-Fallis (1976) refers to parenthetical uses, emphasis, exclamations, repetitions, symmetric alternation, linguistic routines, and anticipation. Other internal factors that Valdés-Fallis points out are random frequent lexical items, lexical need, triggers, stylistic changes, and discourse markers. In turn, Poplack (1980) talks about interjections, idiomatic expressions, tags, and quotes. On the other hand, Gumperz (1992) listed six functions: quotes, listener specification, interjections, reiterations, message qualification, and personalization as opposed to objectivization. Interestingly, Bonvillain, (2008: 321-26) states that, as a bilingual conversational strategy, CS has numerous discourse and interactional functions which she categorizes (cf. Table 3). More recently, Callahan (2004), Jonsson (2010) and Montes-Alcalá (2016) have proved that literary CS is used for a variety of socio-pragmatic and stylistic purposes similar to those found in bilingual speech. All of them propose a number of categories to suit their data; in particular, Montes-Alcalá (2016), following previous classifications, specifically tried to group related functions under the same category, both for clarity and economy’s sake. Table 3 below summarizes the different taxonomies employed by some of these authors, whose categories have partially guided our analysis.

Gumperz (1982)	Bonvillain (2008)	Callahan (2004)	Jonsson (2010)	Montes Alcalá (2016)
quotes	quotes	quotations	quotations	quotations
listener specification	marking discourse boundaries	vocatives	gaps	clarification /elaboration
interjections	to segment phrases marking qualification, explanation or elaboration	expletives	interjection	lexical or cultural need
reiterations	tag questions	referential	reiteration	emphasis
message qualification	attention getter	commentary & repetition	word / language play	stylistic
personalization vs. objectivization	routine phrases	set phrases, tags & exclamations		
	ease of vocabulary / simpler expression	discourse markers		
	to signal changes in syntactic constructions	directives		
	repetition as an emphatic device			
	marking politeness & solidarity			
	shifts within discourse (focus, addressee, topic)			
	interactional & emotional purpose			

Table 3: Classification of the socio-pragmatic functions of CS by several authors

In the following sub-section we will offer our own classification, illustrating the different types of functions that were identified in the samples. Although occasionally there may be some overlapping, the data have been organized according to the most prominent or primary function they seem to perform, taking into account the difficulties that have also been encountered by other researchers in this kind of task. For instance, Montes-Alcalá (2016: 200) warns that “not each and every language switch will always fulfill a unique or specific function”, while Zentella (1997: 99) admits that “no complete

accounting may ever be possible”, in the same way as “imputing the reason for a monolingual’s choice of one synonym over another” is almost impracticable.

4.2.2. Functions of CS in the samples

The following seven functions were identified in the selected novels. I will describe each of them briefly and then a few examples will be provided.

i) *Vocative function*

This is a relatively frequent function that tends to occur whenever Hispanicisms are used by the characters as attention getters, nicknames, address terms or as terms of endearment. In most of the texts studied here Spanish *address terms* abound, with or without a proper noun, for instance *Don Phillip, Señor don Julio, señor, señora, señorita, señorito, Marqués, Marquesa, Madre, Padre, Papa, Abuela, Abuelo, pequeña prima, pequeña hermana, chica, amigo*, etc. The following are some examples:

- (25) ‘This way, *Senorita!*’ (Airlie, 1958: 29)
- (26) ‘In here, *senorito*’ he said. (MacLeod, 1982: 90)
- (27) ‘If you know, *amigo*,’ she asked gently’. (Airlie, 1958: 53)
- (28) ‘You are quite safe with me, *chica*’ (Thorpe, 1973: 11)

Similarly, several *terms of endearment*, such as *amor mío, mi corazón, amada, amada mía, chica mía, querida, mi niña*, have also been found frequently performing this vocative function, as in the utterances below:

- (29) ‘You have made it very plain, *mi niña*, that you want nothing more to do with me.’ (Mayo, 1994: 56).
- (30) ‘There’s nothing to fear in love, *amada*’ (Thorpe, 1973: 112)
- (31) ‘... you must remind me, *querida*.’ (Airlie, 1958: 65)
- (32) ‘You are improving, *chica mía!*’ (Thorpe, 1973: 96)
- (33) ‘*Mi corazón*, I would never do that.’ (Mayo, 1994: 169).
- (34) ‘You are very sweet, *chiquita*,’ he said softly. (MacLeod, 1982: 111)

ii) *Routine formulas*

This discourse function is typically performed by expressions of a sociolinguistic nature, that is, they are related to certain social situations in which they tend to be used recurrently. In fact, the appearance of these routine formulas in discourse involves significant social meaning (Coulmas, 2005: 239). Kecskes (2010) defines them as prefabricated lexical units which are crucial for linguistic fluency. In our sample, they consist mainly of expressions such as *Buenos días, buenas tardes, buenas noches, muchas gracias*, etc. which occur quite frequently in the texts studied. Specific examples follow:

- (35) '*Buenas tardes, señorita.* The *Señor Marqués* awaits you in the *salon*.' (Britt, 1977: 9)
- (36) '*Buenas noches, madrastra!*' (MacLeod, 1982: 58)
- (37) '*Buenos días, señora,*' [...]. '*Cómo está?*' (Lane, 1978: 19)
- (38) '*¡Sí, Señor Martel!*' (Corrie, 1988: 145)
- (39) '*Muchas gracias,*' (MacLeod, 1982: 101)

iii) *Set phrases*

This function is quite similar to the previous one. These set phrases or idiomatic expressions are often inserted in the dialogues. The following are some examples:

- (40) '*Hasta la vista!*' (MacLeod, 1982: 45)
- (41) '*Viva España!*' (Lane, 1978: 103)
- (42) '*Te quiero mucho*' (Airlie, 1958: 50)
- (43) '*Adelante!*' (MacLeod, 1982: 115).
- (44) '*Comprende usted?*' (Mayo, 1992: 118)
- (45) '*¡Salud!*' (Mayo, 1992: 82)
- (46) '*Bienvenida señorita!*' (MacLeod, 1982: 49)

iv) *Quotes*

This is one of the functions that is most frequently performed by the Hispanicisms in the analysed samples; interestingly, all the authors in Table 3 have included it in their classifications. It typically occurs in the texts whenever the narrator opts for directly quoting the exact words uttered by any character, rather than reporting his/her speech:

- (47) The child murmured '*Buenos días.*' (Danbury, 1973: 23)
- (48) Catherine heard Manuel say '*¡Sí, señor!*'. (MacLeod, 1982: 100)
- (49) '*¡Si!*', the boy nodded. (Mayo, 1992: 118)
- (50) '*¡Gracias, Cathy!*' he said. (MacLeod, 1982: 167)
- (51) ...as he came over with a jovial '*Hola*' for his uncle... (Arbor, 1967: 89)
- (52) Alberto [...] greeted Lucie with a '*Buenos días*' (Danbury, 1973: 67)
- (53) '*Buenas noches, señorita!*' she added with a small, mocking laugh. (MacLeod, 1982: 113)

Notice how the Hispanicisms in utterances (48) and (50) above, '*¡Sí, señor!*' and '*¡Gracias!*' obviously perform the function of quoting, but at the same time they may also be categorized as routine formulas. Thus, they are instances of the above-mentioned cases of overlapping.

v) *Interjections*

Interjections are defined as words or phrases that are used to express a strong feeling such as surprise, pain or horror and which are often said loudly and emphatically. With their emphasis on emotions and feelings, interjections are also related to Jakobson's well-known expressive or emotive function, which none of the authors selected in Table 3 mentions. Only Bonvillain (2008) talks about a function that has to do with interactional and emotional purposes, while Callahan (2004) categorizes expletives as another type of function. In the following example the interjection used might also be classified as an expletive:

(54) '*Maldito sea*, Tanya, how can you do this to me?' (Mayo, 1994: 138)

Other cases of Hispanicisms which can be included in this category of interjections because of the emotional sense they involve, are shown below. Notice that (56) can also be a quote:

(55) '*Dios*, woman, I would be better the only one' (Thorpe, 1973: 181)

(56) ...she gave a sudden exclamation. '*Oh dolor!*' (Dansbury, 1973: 92)

vi) *Repetition or reiteration*

This category is employed by almost all the authors in Table 3 above. Apparently, the only exception is Montes-Alcalá (2016), who seems to include this phenomenon in what she calls *emphasis*. In fact, this function occurs when a speaker or character repeats in Spanish what he or she has previously mentioned in English (or viceversa), probably with the aim of explaining the meaning of the Hispanicism or in order to reiterate the message for emphasis. The following examples have been taken from the samples analyzed here:

(57) 'And so *bonito*, so handsome' (Jameson, 1987: 74)

(58) '*Bueno-bueno*, good-good.' (Lane, 1978: 53)

(59) 'Beatriz is my *cuñada*, my sister-in-law' (Mayo, 1994: 60)

(60) 'Nothing, *nada*', she repeated in Spanish'. (Jameson, 1987: 168)

(61) '*Enamorada*, sweetheart, something is wrong?' (Mayo, 1994: 111)

(62) '*Te quiero*. I love you, and I'll marry you!' (Lane 1978: 175)

Interestingly, in utterance (61) the Hispanicism *enamorada* sounds rather artificial; in my opinion, it would have been more natural to use another endearment, such as *cariño*, for instance. Again, this is another case of overlapping as, in addition to the function of repetition or reiteration, this Hispanicism performs the vocative function as well.

(vii) Referential function

It is common knowledge that this is the central function of language. In fact, it is a sort of prerequisite for the rest of functions, as pointed out by Brown (1958: 7), who proves that every single word in a language is right there for some reason and for some purpose, i.e., every word has a referent. This is in line with Page's (1988: 1) idea that "every work of literature possesses in every word that composes it a selective and purposive nature". When assigning this referential function to some cases of CS in her corpus, Callahan (2004: 70-71) noted:

Whether codeswitched material can have a purely referential function is open to interpretation. For example, several of the texts contain Caló vocabulary. It could be argued that the words *jaina* 'woman, girlfriend' and *bato* 'man' communicate certain nuances that their equivalents in standard Spanish would not. However, no terms are entirely neutral; the choice of words from the standard over a regional variety or vice versa communicates still other nuances. But this does not detract from their main function in the discourse: the transmission of information in a narrative.

This referential function can be identified whenever the narrator or any character resorts to a Spanish word to make reference to a particular object or entity. Here a distinction has to be made between those cases in which the choice seems to be justified by a lexical or cultural need, as there is not an exact equivalent in the English language to refer to that object or entity, and those cases in which the writers could have used an English word but opted for the Hispanicism. The former situations are easily explained, as Romaine (1989: 55) states, since "when moving to a new setting, speakers will encounter a variety of things which are specific to the new environment or culture and will adopt readily available words from the local language to describe them." According to Weinreich (1979: 57), the reason for this is merely that "using ready-made designations is more economical than describing things afresh". This lexical or cultural need is simply unavoidable, as the following examples taken from our romances illustrate:

- (63) 'Look, this is a shop where they sell those *timples*', he pointed out. (Danbury, 1973: 115)
- (64) 'Julio has *Guanche* blood.' (Airlie, 1958: 30)
- (65) ...water dripping from his *poncho*,... (MacLeod, 1982: 175)
- (66) 'What shall I play for you, *senoritas*? A *fandango* or a gay *sardana*?' (MacLeod, 1982: 82)
- (67) ...Pasqual cooked *cabrilla* he had caught early that morning (Mayo, 1988: 40)
- (68) Zander explained that this *jameo* was the most famous. (Peake, 1983: 71)
- (69) '... a Canary speciality called *caldo de pescado*, made with a local fish and very tasty...' (Danbury, 1973: 135).
- (70) '...where they dance *flamenco* in the true way'. (MacLeod, 1982: 95)
- (71) ...and she thoroughly enjoyed *Rancho Canario* – a rich vegetable soup... (Mayo, 1988: 116)

- (72) ... to earn a few *pesetas* for their daily bread. (Airlie, 1958: 138)
 (73) ...and balls of *gofio* dipped in a fiery sauce called *mojo*. (Field, 2004: 129)

Very often these switches into Spanish occur only with the insertion of words or phrases which are not followed or preceded by any translation or explanation of their meaning. However, on other occasions the authors immediately provide an English synonym or some kind of clarification of their meaning, as in the sentences below:

- (74) ‘You’ve eaten many kinds of *turrón* - this almond sweet...’ (Danbury, 1973: 135)
 (75) ...the opening parade, the *comparsas* show, the dancers, the fancy dress contest, the songs of the *rondallas*- street musicians – the *murgas* – groups of musical critics... (Mayo, 1994:142)
 (76) ... he veered across a bridge over a *barranco* – a dried up river course (Lane, 1978: 153)
 (77) They’re, I’m told, called *malpais*, that is, modern lava zone. (Peake, 1983: 71)
 (78) ‘*Bacalao*’, confirmed Charlene with a smile, ‘or codfish to you and me, and these - indicating the potatoes, - are *papas arrugadas*, which, translated literally, means wrinkled potatoes...’ (Mayo, 1994: 5)
 (79) ‘*Sí, comprendo!* I understand’, Trudy said quickly. (Lane, 1978: 172)
 (80) This is a special island kind with the cornmeal, *gofio*.(Danbury, 1973: 115)
 (81) *Tapas*, Dervan called them, these tasty bits of food. (Mayo, 1990:41)
 (82) ... when she heard the *silbo*, the unique whistling language native to the island (Mayo, 1988: 199)
 (83) ... not far from the cable-car station was a *parador*, a government-owned hotel ... (Lane, 1978: 56)
 (84) ‘You will not yet have met Madre – my mother.’ (Thorpe, 1973: 26)
 (85) *Zafra* – that is the tomato season-is from late October to the end of April. (Mayo, 1994:102).

Occasionally, explaining the authors’ preference for a switch into Spanish might seem relatively easy, as in the following sentences:

- (86) ‘I have ordered *entremeses* because I thought that Felicity would not care for soup on such a warm evening,’ Sisi informed Philip. (Airlie, 1958: 39)
 (87) ‘I will get my son and take him to see the *eleccion de la reina infantil...*’ (Mayo, 1994:77)

In (86) the author uses the Hispanicism *entremeses*, for which there is not a really English equivalent. The word that is generally used in that language to express the required meaning (‘different dishes of cold food that have been especially prepared to be eaten before the main course of a meal’) originally comes from French, namely, *hors d'oeuvres*.

In turn, although the expression *elección de la reina infantil* in (87) might have relatively easily been translated into “the children’s carnival queen contest”, the author opted for the Hispanicism as it obviously refers to an event which is essentially alien to the Anglosaxon culture.

In contrast, in the sentences below it is not a simple task to explain the reasons why the authors preferred to employ the Spanish terms, instead of their English equivalents, since there is no apparent lexical or cultural need for them:

- (88) ‘My mother was good and sweet and *delicada*.’ (MacLeod, 1982: 93)
- (89) Round the corner strode a *policia* ... (Lane, 1978: 43)
- (90) ‘I need a *siesta*’ Shaine moaned. (Field, 2004: 129)
- (91) ‘We are invited to the *fiesta*,’ she announced. (Airlie, 1958: 69)
- (92) ‘I will show you the *hacienda* at its best.’ (MacLeod, 1982: 52)
- (93) She had drunk too much of Crisógono’s *vino*. (Mayo, 1994: 109)
- (94) ‘...Then I parked her with a large *limonada*, gave Manuel back his red badge of courage...’ (Arbor, 1967: 66)
- (95) ‘I have noticed this *tristeza* in you...’ (Britt, 1977: 54)
- (96) ...the grey ribbons of the *autopista* with heat shimmering from the surfaces (Lane, 1978: 21)
- (97) ‘...on the shaded end of the *patio* outside the *salon* windows...’ (MacLeod, 1982: 66)
- (98) ...as he picked up the *manta* he had discarded on one of the sofas... (Airlie, 1958: 39)
- (99) ‘What do you think of my handsome *sobrino*?’ (Mayo, 1994: 84)
- (100) ‘She is her true *nieta*, full of spirit... (MacLeod, 1982: 62)
- (101) ‘...if your stepmother’s clinic needs another *medico*’. (Lane, 1978: 73)
- (102) ... a long, narrow *barranco*, leading to the sea... (Airlie, 1958: 71)
- (103) ‘Where do courting couples go? Along the *playa*. (Arbor, 1967: 67)

It is interesting to note that a considerable number of the Hispanicisms above are registered in the *Shorter Oxford English Dictionary on Historical Principles* (2007, henceforth *SOEDHP*), that is, they have been officially incorporated as borrowings into the English lexical repertoire, which somehow ‘legitimizes’ their use. In fact, some seem to be relatively well-known by the English-speaking population (such as *fiesta*, *flamenco*, *hacienda*, *macho*, *paella*, *patio*, *sangria*, *señorita*, *siesta*, *tapas*, etc.), while others are not so widely spread, despite their being listed in that same dictionary. It is the case of *barranco*, *bodega*, *cabaña*, *manta*, *meson*, *mirador*, *novio*, *playa*, *silbo*, *tristeza*, *vino*, etc.

The point I want to make here is that many of the Spanish words which are used in the examples above, such as *delicada*, *policia*, *vino*, *limonada*, *tristeza*, *autopista*, *patio*, *salón*, *manta*, *sobrino*, *nieta*, *medico*, *siesta*, *barranco* and *playa*, might well have been replaced respectively by their English equivalents *delicate*, *policeman*, *wine*, *lemonade*, *sadness*, *freeway*, *backyard*, *living-room*, *blanket*, *nephew*, *granddaughter*, *doctor*, *nap*, *ravine* and *beach*. Admittedly, the concepts evoked by some Hispanicisms, such as *fiesta*,

hacienda, *siesta* or *patio* may still add some Spanish flavour to their literal meaning, as long as they highlight some specific cultural features which their English equivalents are felt to lack. Following this line of thought, it could be argued that the idea of an English *backyard*, for example, does not coincide exactly with the image of a Canarian *patio*, which may justify the author's preference to employ this Spanish word that somehow echoes the conceptual and cultural difference in meaning. What seems obvious is that, without all these Hispanicisms, the texts would lose an important expressive effect: the sense of authenticity, realism and local color that they subtly provide with all their connotations, their codified pragmatic value and communicative potential, clearly adding some pragmatic value to the writing (Wotjak, 2006; Khoutyz, 2009; Jiménez Hurtado, 2001).

On the other hand, it has to be acknowledged that the use of Hispanicisms seems to be useful for the authors as a strategy that contributes to making readers aware of the sociolinguistic and sociocultural differences between the characters. It also helps them to construct their identities and to recreate the bilingual context in which the plots are developed. These specific functions cannot be denied and must be mentioned when it comes to describing and explaining the dynamics of the insertion of Hispanicisms in the English discourse of these novels.

5. Conclusion

In this paper I have explored the way Spanish words, phrases and sentences are used in two samples of works taken from a written corpus of English popular romances. The study has offered a two-fold analysis of the Hispanicisms inserted in the 16 novels selected, focusing both on their form and the socio-pragmatic functions they perform. With the examination of the dynamics of these cases of English/Spanish alternation in this doubly stigmatized subgenre, I hope to have contributed to the field of literary CS.

As expected, we have found that in those texts whose protagonists form a mixed (English/Spanish) couple the authors tend to include a higher number of Hispanicisms. Specifically, the analysis of the way these Hispanicisms are incorporated into the English discourse shows the important role they play for the recreation of the bilingual context the novels portray, much more so when the protagonists belong to different cultures and speech communities, forming a mixed couple as is the case of the romances in sample A. In these texts, switches into Spanish seem to be more necessary and thus they tend to be much more frequent as a literary strategy employed by the authors in order to achieve the representation of the bilingual interaction between the lovers. The few exceptions to this rule, namely, Airlie's (1958) and Danbury's (1973) texts in sample B, have been explained on the grounds of the specifications and peculiarities of the heroes and the plots in these two novels, which reveal closer bonds and involvement with the islands.

On the other hand, we have shown that the majority of the Hispanicisms that the authors in the two samples tend to incorporate into their writings are single words (mainly nouns), which means that lexical CS (or Lipski's Type I) predominates. There are also cases of syntactic CS, which include both phrases and sentences, the former being much

more numerous than the latter. These are all examples of Lipski's Type II, which involves inter-sentential CS, while no instances of Type III, or intra-sentential CS, have been found.

In line with the results of other studies, the data in this piece of research reveal that literary CS in these novels tends to perform a number of socio-pragmatic functions which are similar to the ones typically found in bilingual speech. The seven functions identified are the following: vocative function, routine formulas, set phrases, quotes, interjections, repetition or reiteration, and the referential function, with some cases of overlapping between them. Thus, I agree with Callaghan (2004: 2) when she states that "the successful application to a written corpus of a model developed for speech validates the use of written data, and shows that written codeswitching is not inauthentic". Likewise, as already observed, the Hispanicisms which occur in the selected texts oftentimes seem to fulfill some kind of pragmatic marking of their English discourse, by creating associations and meaning extensions. For this reason, I also concur with Montes-Alcalá (2016: 191), who emphasized "the social, pragmatic and cultural nature of CS over the search for grammatical constraints on the phenomenon."

Interestingly, a considerable number of the Hispanicisms employed in the texts studied are already recognized as borrowings as they have been registered in the *SOEDHP* (2007). Therefore, our claim here is that popular romance fiction novels can work as another channel for the spreading of Spanish words and expressions among their wide international readership, which, in turn, may eventually lead to their incorporation as borrowings into the English lexical repertoires. Further research will be needed to confirm this last point; however, we believe that this investigation has been useful to prove that these literary texts can not only successfully depict language contact situations, but they may also function themselves as instruments for language contact and borrowing. Thus, we concur with Braunmüller and House (2009:1-9) that:

for mutual influencing, immediate face-to-face social contact is actually not a necessary condition: lexical borrowing [...] can occur through book reading and learning by writers, teachers, students, who then pass on the new items to other people through other written texts.

Notes

1. Funding received from the Spanish *Ministerio de Economía y Competitividad (MINECO)* for Research Project FFI2014-53962-P, *Discursos, género e identidad en un corpus de novela rosa inglesa ambientada en Canarias y otras islas atlánticas*, is hereby gratefully acknowledged.

2. Following a general tendency in the literature, here we use the terms *switching* and *mixing* as synonyms. However, technically speaking, *code-mixing* refers to special cases of *codeswitching* which are so fast and dense that it is nearly impossible to ascertain which language is being spoken.

3. It is generally believed that Haugen coined the term but, according to Alvarez-Cáccamo (1998), the first scholar to use it was Hans Vogt, in an article published in 1954.

4. Type II is related to the so-called coordinate bilinguals (those who learned their two languages in different times or contexts) and the latter to the compound bilinguals (those who learned both languages at the same time or in the same context). Neither definition is completely satisfactory as most bilingual individuals show features of both types. However, it seems clear that the degree of bilingualism affects the type of CS.

5. Except for the sentences, many of the same Spanish words and phrases appear in several novels; therefore, the corresponding figures in their respective columns cannot be added up.

6. In all the examples given throughout the paper, the Hispanicisms are written in the original form in which they occur in the texts; as shown, sometimes they are neither marked with italics, nor spelt correctly.

References

Primary Sources

Sample A: mixed couples

- Arbor, Jane. *Golden Apple Island*. Harlequin Books, 1967.
 Baird, Jaqueline. *At the Spaniard's Pleasure*. Mills & Boon, 2003.
 Britt, Katrina. *The Villa Faustino*. Mills & Boon, 1977.
 Corrie, Jane. *Island Fiesta*. Mills & Boon, 1980.
 Lane, Pippa. *Nurse in Tenerife*. Mills & Boon, 1978.
 Mayo, Margaret. *Savage Affair*. Harlequin, 1988.
 Mayo, Margaret. *Bitter Memories*. Mills & Boon, 1994.
 McLeod, Jean S. *Meeting in Madrid*. Mills & Boon, 1982.
 Thorpe, Kay. *An Apple in Eden*. Mills & Boon, 1973.
 Wentworth, Sally. *Dark Awakening*. Mills & Boon, 1984.

Sample B: British couples

- Airlie, Catherine. *Red Lotus*. Mills & Boon, 1958.
 Danbury, Iris. *The Silver Stallion*. Mills & Boon, 1973.
 Field, Sandra. *Surrender to Marriage*. Mills & Boon, 2004.
 Jameson, Claudia. *An Engagement is Announced*. Mills & Boon, 1987.
 Mayo, Margaret. *Mutual Attraction*. Mills & Boon, 1990.
 Mayo, Margaret. *Reluctant Hostage*. Mills & Boon, 1992.
 Nickson, Hilda. *Voyage of Discovery*. Harlequin, 1987.
 Oldfield, Elizabeth. *Beware of Married Men*. Mills & Boon, 1986.
 Peake, Lilian. *Night of Possession*. Mills & Boon, 1983.
 Wentworth, Sally. *Fire Island*. Mills & Boon, 1989.

References cited

- Alvarez-Cáccamo, Celso (1988): "From 'switching code' to 'code-switching'. Towards a reconceptualisation of communicative codes". In P. Auer, *Code-switching in Conversation: Language, Interaction and Identity*. London and New York: Routledge, 29-48.
- Anchimbe, Eric A. and Richard W. Janne (2011): "Postcolonial pragmatics: An introduction". *Journal of Pragmatics*, 43(6): 1451-1459.
- Andersen, Gisle, Cristiano Furiassi, and Biljana Misic Ilic (2017): "The pragmatic turn in studies of linguistic borrowing". *Journal of Pragmatics*, 113: 71-76.
- Anderson, Sheri L. (2004): *Status and Solidarity through Codeswitching: Three Plays by Dolores Prida*. MA Thesis. Texas A&M University, College Station, hdl.handle.net/1969.1/536 Accessed 11 March 2017.
- Anderson, Tyler K. and Almeida J. Toribio (2007): "Attitudes towards lexical borrowing and intra-sentential code-switching among Spanish-English bilinguals". *Spanish in Context*, 4(2): 217-240.
- Appel, René and Pieter Muysken (1987): *Language Contact and Bilingualism*. London: Edward Arnold.
- Auer, Peter (1984): *Bilingual Conversation*. Amsterdam: John Benjamins Publishing Company.
- Auer, Peter (1988): *Codeswitching in Conversation. Language, Interaction and Identity*. London and New York: Routledge.
- Bonnin, Juan Eduardo (2014): "Discurso político y cambio de código: la alternancia guaraní / español en el discurso de Fernando Lugo". *Sociolinguistic Studies*, 8(2): 273-291.
- Bonvillain, Nancy. (2008): *Language, Culture and Communication. The Meaning of Messages*. New Jersey: Prentice Hall.
- Brown, Roger (1958): *Words and Things. An Introduction to Language*. Illinois: The Free Press/The MacMillan Company.
- Braunmüller, Kurt and Julianne House (2009): *Convergence and Divergence in Language Contact Situations*. Amsterdam: John Benjamins Publishing Company.
- Bullock, Barbara and Almeida Jaqueline Toribio (2009): *The Cambridge Handbook of Linguistic Code-Switching*. Cambridge: Cambridge University Press.
- Callahan, Laura (2004): *Spanish/English Codeswitching in a Written Corpus*. Amsterdam: John Benjamins Publishing Company.
- Clyne, Michael (2003): *Dynamics of Language Contact: English and Immigrant Languages*. Cambridge: Cambridge University Press.
- Cortés-Conde, Florencia and Diana Boxer (2002): "Bilingual word-play in literary discourse: The creation of relational identity." *Language and Literature*, 11(2): 137-151.
- Coulmas, Florian (2005): *Sociolinguistics: The Study of Speakers' Choices*. Cambridge: Cambridge University Press.
- Culpeper, Jonathan (2014): *Language and Characterization. People in Plays and Other Texts*. Hoboken: Taylor and Francis.
- Curell, Clara et al. (2009): "Por un glosario de hispanismos en la Literatura de Viajes". In C. Solé i Castells, M. Parra, P. Solá, M.C. Figuerola, eds., *Texto y sociedad en las letras francesas y francófonas*. Universitat de Lleida, 634-641.
<http://dialnet.unirioja.es/servlet/oaiart?codigo=3711598>. [Accessed 21 Jan. 2017].
- Curell, Clara and Cristina G. de Uriarte (2001): "Hispanismos en los libros de viaje franceses del siglo XVIII". In I. Uzcanga, E. Llamas and J.M. Pérez, eds., *Presencia y renovación de la lingüística francesa*. Salamanca: Ediciones Universidad de Salamanca, 99-107.

- Eastman, Carol M. (1992): *Codeswitching*. Clevedon: Multilingual Matters.
- Flores Ohlson, Linda (2008): 'Soy el brother de dos lenguas.' *El cambio de código en la música popular contemporánea*. Acta Universitatis Gothoburgensis. Göteborgs Universitet. Romanica Gothoburgensia 55.
- Gardner-Chloros, Penelope (1991): *Language Selection and Switching in Strasbourg*. Oxford: Clarendon Press.
- Gardner-Chloros, Penelope (2015): "Code switching – where next?" Paper presented for a Plenary Session at the *International Conference on Bilingualism*, held at the University of Malta, 23-25 March.
- Giles, Howard and Smith, Philip (1979): "Accommodation theory: Optimal levels of convergence". In H. Giles and R. N. St. Clair, eds., *Language and Social Psychology*. Oxford: Basil Blackwell, 45–65.
- Giménez-Eguibar, Patricia and De los Heros, Susana (2016): "La alternancia lingüística en los diálogos de *East Los Angeles High*. Una forma de representación mediática con función afiliativa de la latinidad". In D. Dumitrescu and D. Bravo, eds., *Roles situacionales, interculturalidad y multiculturalidad en encuentros en español*. Buenos Aires: Editorial Dunken, 31-54.
- González-Cruz, María-Isabel (1995): *La convivencia anglo-canaria. Estudio sociocultural y lingüístico, 1880-1914*. Las Palmas de Gran Canaria: Cabildo Insular de Gran Canaria.
- González-Cruz, María-Isabel (2002): *Notas para una bibliografía inglesa sobre Canarias*. La Laguna: Instituto de Estudios Canarios / Universidad de La Laguna.
- González-Cruz, María-Isabel (2011a): "Hispanismos y canarismos en los textos de dos viajeras inglesas decimonónicas". *Revista de Filología de la Universidad de la Laguna*, 29: 81-99.
- González-Cruz, María-Isabel (2011b): "Exploring the Canarian contribution to Hispanicisms in English". *Revista Alicantina de Estudios Ingleses*, 24: 131-152.
- González-Cruz, María-Isabel (2013): "Hispanismos y canarismos en un corpus de textos ingleses sobre Canarias." *LEXIS. Revista de Lingüística y Literatura*, 37(2): 229-267.
- González-Cruz, María-Isabel (2014): "Nuevos registros para la bibliografía en lengua inglesa sobre Canarias: siete relatos de viaje, dos guías turísticas, una obra de ficción y un estudio lingüístico." *Anuario de Estudios Atlánticos*, 60: 599-641.
- González-Cruz, María-Isabel (2015): "Love in Paradise: Visions of the Canaries in a corpus of popular romance fiction novels". *Oceanide. Journal of the Spanish Society for the Study of Popular Culture SELICUP7*, <http://oceanide.netne.net/articulos/art7-4.pdf>
- González-Cruz, María-Isabel (2017): "Conciencia sociolingüística e hispanismos en un corpus de novela rosa inglesa." *SOPRAG Pragmática sociocultural / Sociocultural Pragmatics*, 5(2): 125-149. <https://doi.org/10.1515/soprag-2017-0014>
- González Cruz, M^a Isabel and M^a del Pilar González de la Rosa (2006): "Language and travel: Spanish vocabulary in British travel books." *Tesserae. Journal of Iberian and Latin American Studies*, 12(2-3): 203-217.
- González Cruz, M^a Isabel and M^a del Pilar González de la Rosa (2007): El viaje lingüístico: una aproximación sociolingüística a la literatura de viajes". In J. M. Oliver, C. Curell, C. G. Uriarte and B. Pico, eds., *Escrituras y reescrituras del viaje. Miradas plurales a través del tiempo y de las culturas*. Bern: Peter Lang, 235-251.
- Grosjean, François (1982): *Life with Two Languages. An Introduction to Bilingualism*. Cambridge (Massachusetts): Harvard University Press.
- Gumperz, John Joseph (1992): *Discourse Strategies*. Cambridge: Cambridge University Press.

- Haugen, Einar (1956): "Bilingualism in the Americas: A bibliography and research guide". *Publications of the American Dialect Society* 26. University of Alabama Press.
- Heller, Monica, ed. (1988): *Codeswitching: Antropological and Sociolinguistic Perspectives*. New York: Mouton de Gruyter.
- Heller, Monica (1992): "The politics of codeswitching and language choice". *Journal of Multilingual and Multicultural Development*, 13(1-2): 123-142.
- Heredia, Roberto R. and Jeanette Altarriba (2001): "Bilingual language mixing: Why do bilinguals code-switch?" *Current Directions in Psychological Science*, 10(5): 164-172.
- Jacobson, Rodolfo, ed. (1998): *Codeswitching Worldwide*. Berlin and New York: Mouton de Gruyter.
- Jacobson, Rodolfo (2001): *Codeswitching Worldwide II*. Berlin and New York: Mouton de Gruyter.
- Jackson, Jane (2014): *Introducing Language and Intercultural Communication*. London and New York: Routledge.
- Jiménez Hurtado, Catalina (2011): *Léxico y pragmática*. Bern: Peter Lang.
- Jonsson, Carla (2010): "Functions of code-switching in bilingual theater: An analysis of three Chicano plays". *Journal of Pragmatics*, 42(5): 1296-1310.
- Kanellos, Nicolás (2003): *Hispanic Literature of the United States: A Comprehensive Reference*. Westport, CT: Greenwood Publishing.
- Keller, Gary (1984): "How Chicano authors use bilingual techniques for literary effect". In E. García et al., eds., *Chicano Studies: A Multidisciplinary Approach*. New York: Teachers College, 171-192.
- Kecskes, Istvan (2010): "Formulaic language in English Lingua Franca". In P. Hanks and R. Giora, eds., *Metaphor and Figurative Language: Critical Concepts in Linguistics*. London and New York: Routledge.
- Khoutyz, Irina (2009): "Anglicisms as a means of pragmatic marking". *Scientific Bulletin of Politehnica University of Timisoara*, 8(1-2): 5-11.
- Lipski, John M. (1982): "Spanish English language switching in speech and literature – theories and models". *The Bilingual Review*, 9(3): 191-212.
- Lipski, John M. (2005): "Codeswitching or Borrowing? No sé so no puedo decir, you know". In L. Sayahi and M. Westmoreland, eds., *Selected Proceedings of the Second Workshop on Spanish Sociolinguistics*. Somerville, MA: Cascadilla Press, 1-15.
- L. Sayahi and M. Westmoreland (2014): "Spanish-English codeswitching among low-fluency bilinguals: Towards an expanded Typology". *Sociolinguistic Studies*, 8:23-55.
- Milroy, Lesley and Pieter Muysken, eds. (1995): *One Speaker, Two Languages. Cross-Disciplinary Perspectives on Code-Switching*. Cambridge: Cambridge University Press.
- Montes-Alcalá, Cecilia (2000a): "Written code-switching: Powerful bilingual images". In R. Jacobson, *Codeswitching Worldwide II*. Berlin and New York: Mouton de Gruyter, 59-74.
- Montes-Alcalá, Cecilia (2000b): "Attitudes towards oral and written code-switching in Spanish-English bilingual youth". In A. Roca, ed., *Research on Spanish in the United States: Linguistic Issues and Challenges*. Somerville, MA: Cascadilla Press, 218-227.
- Montes-Alcalá, Cecilia (2001): "Oral vs. written code-switching contexts in English-Spanish bilingual narratives". In I. de la Cruz, ed., *La lingüística aplicada a finales del siglo XX. Ensayos y propuestas*, vol. 2. Madrid: Servicio de Publicaciones de la Universidad de Alcalá de Henares, 715-720.
- Montes-Alcalá, Cecilia (2005a): *Two Languages, One Pen: Sociopragmatic Functions in Written Spanish-English Codeswitching*. PhD Thesis. Michigan: Hell & Howell Information.

- Montes-Alcalá, Cecilia (2005b): "Mándame un e-mail: cambio de códigos español-inglés online". In L. Ortiz López and M. Lacorte, eds., *Contactos y contextos lingüísticos: El español en los Estados Unidos y en contacto con otras lenguas*. Madrid and Frankfurt: Iberoamericana, 173-185.
- Montes-Alcalá, Cecilia (2012): "Code-switching in US Latino novels". In M. Sebba, S. Mahootian and C. Jonsson, eds., *Language Mixing and Code-Switching in Writing: Approaches to Mixed-Language Written Discourse*. London and New York: Routledge, 68-88.
- Montes-Alcalá, Cecilia (2015): "Code-switching in US Latino Literature: The role of biculturalism". *Language and Literature*, 24(3): 264-281.
- Montes-Alcalá, Cecilia (2016): "Socio-pragmatic functions of codeswitching in Nuyorican and Cuban American literature". In G. Tamargo et al., eds., *Spanish-English Codeswitching in the Caribbean and the US*. Amsterdam: John Benjamins Publishing Company, 191-213.
- Montes Granado, Consuelo (2012): "Code-Switching as a strategy of brevity in Sandra Cisneros' *Woman Hollering Creek* and other stories". *DQR Studies in Literature*, 49: 125-138.
- Moyer, Melisa (1998): "Bilingual conversation strategies in Gibraltar". In P. Auer, ed., *Codeswitching in Conversation*. London and New York: Routledge, 215-234.
- Müller, Katharina B. (2015): "Code-switching in Italo-Brazilian literature from Rio Grande do Sul and Sao Paulo: A sociolinguistic analysis of the forms and functions of literary code-switching". *Language and Literature*, 24(3): 249-263.
- Muysken, Pieter (2000): *Bilingual Speech*. Cambridge: Cambridge University Press.
- Myers-Scotton, Carol (1992): "Comparing codeswitching and borrowing." *Journal of Multilingual and Multicultural Development*, 13(1-2): 19-39.
- Myers-Scotton, Carol (1993): *Social Motivations for Codeswitching. Evidence from Africa*. Oxford: Clarendon Press.
- Page, Norman (1988): *Speech in the English Novel*. London: MacMillan.
- Pennycook, Alistair (2012): *Language and Mobility: Unexpected Places*. Bristol: Multilingual Matters.
- Pfaff, Carol W. and Laura Chávez (1986): "Spanish/English codeswitching: Literary reflections of natural discourse". In R. von Bardeleben et al., eds., *Missions in Conflict: Essays on U.S.-Mexican Relations and Chicano Culture*. Mainz: Gunter Narr Verlag Tubingen, 229-254.
- Poplack, Shana (1980): "Sometimes I'll start a sentence in Spanish Y TERMINO EN ESPAÑOL: Toward a typology of code-switching". *Linguistics*, 18(7-8): 581-618.
- Poplack, Shana and Marjory Meechan (1998): "How languages fit together in codemixing". *International Journal of Bilingualism*, 2: 127-138.
- Romaine, Suzanne (1995): *Bilingualism*. Oxford: Basil Blackwell.
- Romance Writers of America n.d. "About the romance genre". Available at <http://www.rwa.org/p/cm/ld/fid=578> [Accessed 19 March 2016].
- Sánchez-Palencia Carazo, Carolina (1997): *El discurso femenino de la novela rosa en lengua inglesa*. Cádiz: Publicaciones de la Universidad de Cádiz.
- Sebba, Mark, Shahrzad Mahootian, and Carla Jonsson, eds. (2012): *Language Mixing and Code-Switching in Writing: Approaches to Mixed-Language Written Discourse*. London and New York: Routledge.
- Sebba, Mark (2000): "Writing switching in British Creole". In K. Jones and M. Martin-Jones, eds., *Multilingual Literacies: Reading and Writing Different Worlds*. Amsterdam: John Benjamins Publishing Company, 171-187.
- Shorter Oxford English Dictionary on Historical Principles* (2007): Oxford: Oxford University Press.

- Söderqvist, Therése (2009): 'All of a sudden me siento tristoná.' *El cambio de código inglés / español en MySpace*. Unpublished MA Thesis. Göteborgs Universitet, Available at https://gupea.ub.gu.se/bitstream/2077/20834/1/gupea_2077_20834_1.pdf. [Accessed 10 May 2016].
- Tabouret-Keller, Andrée (1997): "Language and identity". In F. Coulmas, ed., *The Handbook of Sociolinguistics*. Malden, USA and Oxford, UK: Wiley-Blackwell, 315-326.
- Thompson, Gregory L. and Edwin M. Lamboy (2012): *Spanish in Bilingual and Multilingual Settings Around the World*. Leiden and Boston: Brill.
- Valdés-Fallis, Guadalupe (1976): "Code-switching in bilingual Chicano poetry". *Hispania*, 59: 866-877.
- Valdés-Fallis, Guadalupe (1977): "The sociolinguistics of Chicano literature: Towards an analysis of the role and function of language alternation in contemporary bilingual poetry". *Punto de contacto*, 1(4): 30-39.
- Weinreich, Uriel (1979): *Languages in Contact. Findings and Problems*. New York: Mouton.
- Woolard, K.A. (1987): "Code-switching and comedy in Catalonia". *International Pragmatics Association Papers in Pragmatics*, 1: 106-122.
- Wotjak, Gerd (2006): *Las lenguas, ventanas que dan al mundo. El léxico como encrucijada ente morfosintaxis y cognición: aspectos semánticos y pragmáticos en perspectiva intra e interlingüística*. Salamanca: Ediciones Universidad de Salamanca.
- Zentella, Ana Celia (1997): *Growing up Bilingual. Puerto Rican Children in New York*. Oxford: Basil Blackwell.

Alicante Journal of English Studies / Revista Alicantina de Estudios Ingleses (hereinafter, *RAEI*) is published yearly by the Department of English at the University of Alicante in volumes of approximately 250 pages. The Journal aims to provide a forum for debate and an outlet for research involving all aspects of English Studies.

NATURE AND FORMAT OF THE ARTICLES: The *Revista* would welcome articles of the following kinds: (1) Full-length articles (8,000-12,000 words in length) on all topics within the field of English studies (linguistics, literary criticism, literature, teaching methodology, and translation and interpreting); (2) Reviews (around 1,500-2,000 words) of recently published books (appeared within two years from the date of submission). All articles submitted should follow the guidelines which can be obtained from the following Internet address:

<http://raei.ua.es/index>

Manuscripts should include an abstract in English of about one hundred words in length and five keywords. In normal circumstances, the editors will only consider for publication those contributions written in English and in MS Word format. All correspondence should be addressed to raei@ua.es, or, alternatively, to:

Revista Alicantina de Estudios Ingleses, Departamento de Filología Inglesa, Universidad de Alicante, P.O. Box 99, E-03080 ALICANTE (Spain).

RAEI is indexed in the following citation databases: (1) MLA International Bibliography; (2) Linguistics Abstracts; (3) Languages and Literatures; (4) Periodicals Index Online; (5) Summarev (journal summaries from the University of Seville); (6) Dialnet; (7) ISOC (CSIC database); (8) Bibliographie Linguistique (BL Online).

- **ADVERTISING:** The journal will be pleased to carry advertisements in either full-page (17 x 24 cms. approx.) or half-page (17 x 12 cms. approx.) format. Prices and information are available on request at the above address.
- **EXCHANGES:** *RAEI* will be happy to make exchange arrangements with similar journals in the same field. All such proposals should be made in writing to the above address.
- **SUBSCRIPTIONS:** The price for the annual subscription to *Revista Alicantina de Estudios Ingleses* (one volume per year) is as follows: (1) in Spain, 30 euros for libraries and institutions, and 20 euros for private subscribers; (2) in countries other than Spain, US \$35 for libraries and institutions, and US \$28 for private subscribers.

Correspondence on subscriptions should be addressed to:

Marcial Pons Librero
San Sotero 6.
28037 MADRID (Spain)
revistas@marcialpons.es
Tel: +34 913043303
Fax: +34 913272367

Alicante Journal of English Studies

Revista Alicantina de Estudios Ingleses

n° 30, December 2017

INTRODUCTION: Gaining insights into English as a contact language and its diffusion, by Ignacio M. Palacios Martínez and Jose A. Sánchez Fajardo.

ESSAYS: English as a Lingua Franca (ELF) in ESP contexts. Students' attitudes towards non-native speech and analysis of teaching materials, by Yolanda Joy Calvo-Benzies; Rhoticity in Chinese English: An experimental investigation on the realization of the variant (r) in an Expanding Circle variety, by Zeyu Li and Baris Kabak; Language contact in Gibraltar English: A pilot study with ICE-GBR, by Lucía Loureiro-Porto and Cristina Suárez-Gómez; Contact-induced variation in clausal verb complementation: The case of *regret* in World Englishes, by Raquel P. Romasanta; Syntactic complexity and language contact: A corpus-based study of relative clauses in British English and Indian English, by Iván Tamaredo; English lexical items in Egyptian Arabic. Code-switching or borrowing?, by Malgorzata Kniaz; How *VIPs* and *IT-professionals* behave in Russian. The use of loan abbreviations in Russian analytical composites, by Elizaveta Tarasova; Present-day Spanish fashion lexicon dresses up in English, by Beatriz Rodríguez Arrizabalaga; Analysis of the presence of anglicisms in a Spanish internet forum: Some terms from the fields of fashion, beauty and leisure, by Carmen Luján García; The influence of the English language on the description of cosmetic products, by Cristina Tejedor Martínez; Exploring the dynamics of English/Spanish codeswitching in a written corpus, by María Isabel González Cruz.