Terminology in tourism 2.0: identification of the categories in user generated reviews

María Luisa Carrió-Pastor
Universitat Politècnica de València
Camino de Vera, s/n
46022 Valencia
E-mail: lcarrio@upv.es

Miguel Ángel Candel Mora
Universitat Politècnica de València
Camino de Vera, s/n
46022 Valencia
E-mail: mcandel@upv.es

TERMINOLOGY IN TOURISM 2.0: IDENTIFICATION OF THE CATEGORIES IN USER GENERATED REVIEWS

TERMINOLOGÍA EN TURISMO 2.0: IDENTIFICACIÓN DE CATEGORÍAS EN LAS CRÍTICAS GENERADAS POR LOS USUARIOS

LA TERMINOLOGIE DU TOURISME 2.0: L'IDENTIFICATION DES CATÉGORIES DANS LES AVIS DES UTILISATEURS

ABSTRACT: This paper focuses on the analysis of consumer-generated reviews in travel websites and their use of terminology. The objectives of this paper are first, to study the categories most commonly used in consumer reviews, second, to analyse the use of borrowings and neologisms in consumer reviews and finally, to identify the different words used by reviewers to refer to the same concept and the reason for this variation. In order to reach these objectives, a corpus of two hundred consumer reviews written in Spanish by users of an online platform was compiled. On the one hand, the terms were identified and classified attending to semantic criteria and their level of specialisation. On the other, the borrowings and neologisms were identified in the corpus, being selected and classified in order to study the variation of terms. Then, the results were extracted and discussed and finally, conclusions were drawn.

RESUMEN: Este artículo se centra en el análisis de las opiniones generadas por los usuarios en sitios web de viajes y su uso de la terminología. Los objetivos de este trabajo son, en primer lugar, estudiar las categorías de términos más utilizadas en las opiniones de los usuarios; en segundo lugar, analizar el uso de préstamos y neologismos y, finalmente, identificar las diferentes palabras utilizadas por los usuarios para referirse al mismo concepto y la razón de esta variación. Para alcanzar estos objetivos, se recopiló un corpus de doscientas opiniones escritas en español por usuarios de una plataforma en línea. Por un lado, se procedió a la identificación de términos y posteriormente se clasificaron atendiendo a criterios semánticos y su nivel de especialización. Por otro lado, se identificaron los préstamos y neologismos del corpus, y se seleccionaron y clasificaron para observar la variación de los términos. Posteriormente, se ofrece una discusión de los resultaros y finalmente, se presentan las conclusiones.

RÉSUMÉ: Cet article est centré sur l'analyse des avis des utilisateurs des sites de voyages et leur utilisation de la terminologie. Les objectifs du travail sont, en premier lieu, l'étude des catégories des termes les plus employés dans les avis des utilisateurs ; en second lieu, l'analyse de l'utilisation d'emprunts et de néologismes et, enfin, l'identification des mots divers utilisés par les utilisateurs pour faire référence à un même concept ainsi que la raison de cette variation. Afin d'atteindre ces objectifs nous avons créé un corpus de deux cents avis écrits en espagnol par des utilisateurs d'un portail en ligne. D'un côté, nous avons procédé à l'identification des termes, par la suite ils ont été classés selon des critères sémantiques et leur niveau de spécialisation. De l'autre côté, les emprunts et les néologismes ont été identifiés dans le corpus, puis ils ont été sélectionnés et classés afin d'observer la variation des termes. Une discussion des résultats est proposée ensuite en enfin les conclusions sont tirées.

KEY WORDS: web 2.0; neologisms; consumer reviews; tourism; borrowings.

PALABRAS CLAVES: web 2.0; neologismos; opiniones de los usuarios; turismo; préstamos.

MOTS CLÉS: web 2.0; néologismes; avis des utilisateurs; tourisme; emprunts.

SUMMARY: 1. Introduction. 2. Corpus and method. 3. Analysis and discussion of results. 4. Conclusions.

SUMARIO: 1. Introducción. 2. Corpus y método. 3. Análisis y discusión de los resultados. 4. Conclusiones.

SOMMAIRE: 1. Introduction. 2. Corpus et méthodologie. 3. Analyse et discussion des résultats. 4. Conclusions.

Terminology in tourism 2.0: identification of the categories in user generated reviews

María Luisa Carrió-Pastor & Miguel Ángel Candel Mora

1. Introduction

This paper draws on three linguistic aspects. The first concerns the patterns with which to identify and classify specific words (Dubuc and Lauriston, 1997, Lerat, 1997, Sager, 1997, Resche, 2000, Alcaraz, 2000, Carrió Pastor and Candel Mora, 2013); the second, the identification of neologisms and their classification (Cabré, 2002, 2006, Gotti, 2003); and finally, the importance of digital literacies in communication (Scollon and Scollon, 2004, Skovholt and Svennevig, 2013, Darics, 2015).

As the definition of terminology has attracted the interest of many researchers since Wüster (1974) initiated this field of study (Sager, 1990, 1999, Cabré, 1999, 2003, Temmerman, 2000), we will not focus on this here. Rather, we will concern ourselves with the level of specialization of terms. According to Alcaraz (2000: 15), "(…) a specialized language is a specific language used by professionals and specialists to transmit information and to negotiate terms, concepts and knowledge in a specific area of knowledge". The main features that characterize specialized communication have also been classified into three categories by Cabré (1998: 29): the specificity of the subject matter from a cognitive perspective; the interlocutors, since users are generally specialists in a particular field, and terminology, as the higher the degree of specialization of a text, the higher its terminological density.

Furthermore, the lexical aspects that characterize specialized texts make it possible different levels of specialization to be distinguished and a classification to be made, consisting of technical terms, semi-technical terms, neologisms, and compounds and derivatives, following the guidelines proposed by Dubuc and Lauriston (1997), Cabré (1998) and Alcaraz (2000). In this study, we follow the above classification system, which holds that technical terms are those traditionally belonging to science and technology and which are recorded in technical dictionaries, glossaries or scientific and technical texts. Semi-technical terms make up the greater part of the specialized language of any discipline, since their use is not limited to scientific and technical texts: they come from the general language but designate different concepts depending on the context and the specialized field concerned. Neologisms are those terms which have arisen with the growth of scientific and technical areas but are not yet recorded in dictionaries. Finally, compounds and derivatives emerge from word formation processes. Today, the Internet is changing the way we communicate and thus new terms are being incorporated into well-known classifications of terminology due to the emergence

of new genres. We believe this is an important aspect to take into account when analysing terms.

There are also categories of tourism terminology that follow semantic criteria that are of interest, such as the one proposed by the Secretariat of State for Tourism of France and the World Tourism Organization (2001):

01. Sports
02. Tourism Legislation
03. Ecology of Tourism
04. Economy of Tourism
05. Tourism Facilities
06. Visitor Flows
07. Training and Employment
08. Accommodation
09. Leisure Activities
10. Tourism Events
11. Tourism Heritage
12. Tourism Policy
13. Tourism Services
14. Tourism Professionals
15. Tourism Promotion
16. Science and Information
17. Sociology of Leisure
18. Tourism Sectors
19. Transport
20. Countries and Country Groupings

Also, Alcaraz, Hughes, Campos, Pina and Aleson (2000) proposed a shorter classification of terms in the semantic fields of tourism and leisure:

1. Accommodation
2. Art/culture
3. Catering
4. Climate
5. Entertainment
6. Insurance
7. Landscape
8. Management
9. Marketing
10. Sports
11. Travel
12. General
13. Games

These two specific classifications of terms that follow semantic criteria have been very useful for our research and this paper also proposes an ad hoc classification of eight categories based on the results extracted from the analysis of the corpus.

In this study, we will also focus on neologisms, as most of them are part of specialized texts: more precisely, we will examine the neologisms used in tourism 2.0. Newmark (1988) defined neologisms as newly coined lexical units and Cabré (1999) also remarked that "(...) a unit is a neologism if it has arisen recently, if it is not in dictionaries, if it exhibits signs of formal instability (...) or semantic instability and if speakers perceive it as a new unit". More recently, Kessler (2010: 262) has also described neologisms as being part of technology in the sense that "(...) technology oriented neologisms are often too ephemeral and numerous for even trained observers to adequately document". At this juncture, it can be said that neologisms are being created every day, due to the emergence of new ways of communication, and this is the reason why a large number of researchers are now turning their attention to this field of study (Cabré and Estopà, 2004, Cabré, 2006, Worsoe, 2011, Talebinejad, Dastjerdi and Mahmoodi, 2012, Cabré, Estopà and Vargas, 2012, Moghadam and Sedighi, 2012).

In this paper, the focus is on the identification and classification of neologisms, specifically in the field of tourism and in the genre of online consumer reviews. In recent years, a range of proposals has been made for the classification of these novel terms. Cabré and Estopà (2004) established the following classification: formal neologisms, syntactic neologisms, semantic neologisms, borrowings and other kinds of neologisms (including cases which are difficult to classify such as jargon, dialectal words, etc.). Later, Cabré (2006) discussed the difficulty of classifying neologisms and refined the categories established previously, including the category of variation. The classification proposed by this author can be seen in Table 1:

Formal neologisms	Variation	Syntactical neologisms	Semantic neologisms	Borrowings	Other
Suffixes	Orthographic variation	Change of grammatical category	Change in a lexical base	Orthographically adapted borrowings	Dialectal words
Prefixes			Proper name	Non-adapted	Slang
Composition					Jargon
Lexicalization					
Syntactic change					
Lexicalization of a syntactic structure					
Acronyms					
Abbreviation					

Table 1: Classification of neologisms proposed by Cabré (2006)

Another interesting proposal was made by Megerdoomian and Hadjarian (2010), who analysed neologisms in blogs. These authors proposed the classification of the formation of neologisms into borrowing, compounding, affixation and conversion (semantic and functional shifts). More recently, Barbu (2011) proposed a more descriptive classification that focuses on the neologisms created in Romance languages. She distinguished the following types of neologisms: derivation with affixes, compounds, elevated compounds, conversion, syntactic neologisms, graphical and morphological doublets, English loan words, idiosyncratic loan words and others (dialectal words, slang, professional jargon). Also, Cabré and Nazar (2011) proposed a simpler classification: formal neologisms (monolexical and polylexical) and semantic neologisms. The neologisms found in the corpus analysed here were classified according to the proposal made by Cabré (2006), as it is more detailed and the neologisms can thus be more easily differentiated.

Finally, the issue of digital literacies (Scollon and Scollon, 2004, Skovholt and Svennevig, 2013, Carrió Pastor 2015) also forms an important part of this study. Digital communication has become increasingly important for the everyday exchange of information (Crystal 2001). Internet users communicate with people all around the world in just a few seconds and the near-instantaneous nature of this has taken on great significance in some professional environments, such as tourism. In this paper, our specific interest lies in the use of the Internet in the provision of opinions in the form of consumer reviews. Taking into account the communication and exchange of ideas that now occurs on the Web, Web 1.0 users were effectively spectators and in Web 2.0 they have become active participants, since Web 2.0 is characterized by its power to create networks and build relationships around common interests, opinions and product information. This is the reason why it has gained massive popularity (O'Reilly, 2007, Cox, et al. 2008).

The impact of technology on tourism has been such that we can find evidence in Tourism 2.0 of the social and cultural conventions of speakers (Edo, 2012, Schemmann, 2011). Although knowledge and its transfer should be the motor of these networks that self-organise and develop autonomously on the basis of the contributions of their members, online consumers have different educational, cultural and linguistic backgrounds. This may create an abundance of variation in digital communication and we believe that this specialised lexis is worthy of analysis, and a further analysis of the different linguistic codes could then be undertaken (Goethals, 2014).

Undoubtedly, consumer-generated reviews now play a fundamental role in the tourism sector as Web 2.0 has enabled the active participation of users to such an extent that, in the field of travel and tourism services, reviews have become a much more reliable source of information than the service provider's own marketing campaigns. At first sight, several factors may have an effect on the language used in consumer reviews, such as anonymity, the

reviewer's experience, how easy it is to post a review on the platform, the medium used (PC, tablet, mobile phone, etc.), or even what is being reviewed, which may be a restaurant, a hotel, flights, a tourist attraction, and so on (Schemmann, 2011).

Consumer-generated content comprises a variety of forms and types. According to Schemmann (2011), consumer reviews in the tourism context fall into three categories, (1) service evaluation, (2) feedback and interactive functions and (3) matching and search performance functions. The first category includes free-style text and structured text, typically reviews which vary in style and length, and with consumers also being able to recommend a certain service. The other two functions, less frequently found on online platforms, involve feedback, interactive functions and the rating of services. For this paper, the corpus was composed of free-style texts because the objective was to identify the role of specialisation and investigate how this new medium handles the use of new concepts in the form of neologisms.

Consumer reviews enable two-way communication between the tourist industry and consumers. As Vásquez (2014: 2) explains, "(…) the online review seems to have rapidly evolved into a communicative genre that many of us have come to take for granted (…) we rarely stop to think about the numerous linguistic choices that were involved in the actual construction of those texts". In this sense, we believe that consumer reviews are a new genre (Calvi, 2010, Seargeant and Tagg, 2014) with terminology and neologisms that need to be identified and classified in order for the evolution of tourism 2.0 to be understood.

The objectives of this paper are first, to study the categories more commonly used in consumer reviews, second, to analyse the use of borrowings and neologisms in consumer reviews and finally, to identify the different words used by consumer reviewers to refer to the same concept and the reason of this variety.

2. CORPUS AND METHOD

First, two hundred user reviews written in Spanish by Spanish customers on the online platform TripAdvisor (www.tripadvisor.com) were collected taking into consideration they were written by Spanish speakers and the length of the comments, i.e. they should be longer than 3 words. This corpus was then analysed with *WordSmith Tools 5.0* to accurately calculate the overall statistics, extract frequency wordlists and, later, analyse the occurrences of the items selected for this study in context. Also, the terms and neologisms were extracted from the data using the *SDL Multiterm Extract software*. Although this tool was originally intended to be used along with a computer-assisted translation software package, it can handle large monolingual corpora and thus, the automatic extraction process used made the initial task of selecting terms from the corpus much more straightforward. The application's extraction parameters were adjusted to standard noise extraction

since all the term candidates had to be hand edited; the maximum word length was set to 1 element to avoid the extraction of compounds, which might increase noise in the initial extraction; and, finally, a stopword list was added.

The terms were identified and classified attending to their semantic criteria, first with the tools described and afterwards manually. The neologisms were also identified manually and classified into categories. As the only kind of neologisms found were borrowings, they were classified depending on whether they were borrowings from English or not. As borrowings from other languages were not detected, the borrowings from English were also classified. It was decided to group them according to their semantic categories to determine whether the topic of the review had an effect on the use of terminology.

When the results were extracted, quantitative and qualitative analyses were carried out. Finally, the characteristics of the terms and neologisms were identified, the classifications made and the conclusions drawn.

3. ANALYSIS AND DISCUSSION OF RESULTS

The descriptive data obtained from the analysis of the corpus with Word-Smith Tools can be seen in Table 2:

Categories	Raw data of the corpus
Running words	15,451
Tokens used for word list	15,206
Types (distinct words)	2,733
Type/token ratio (TTR)	17.97
Standardised TTR	41.76
Mean word length	4.66
Sentences	859
Mean (in words)	17.70

Table 2: Basic statistics of the corpus of consumer reviews

The two hundred consumer reviews were composed of 15,451 running words, and the distinct items detected by the tool amounted to a total of 2,733. Once the initial analysis of the corpus had been performed, the data related to the terms and their classification was analysed.

3.1. ANALYSIS OF THE RESULTS

After performing the automatic terminology extraction process and subsequent manual examination of the findings, the terms found in the corpus were classified into nine broad categories or semantic fields of tourism shown in Table 3:

Categories of semantic fields of tourism
1. Facilities

2.	Room
3.	Management
4.	Review platform
5.	General
6.	Catering
7.	Staff
8.	Leisure activities
9.	Technology

Table 3: Categories of terms related to tourism

After the first extraction with MultiTerm Extract, 275 terms were identified, of which 34 were validated as being directly related to the terminology of tourism, the object of this study. In addition, another 28 terms were extracted manually as they were not identified by MultiTerm Extract's automatic extraction process.

Thus, overall, 62 terms appeared a total of 801 times in the corpus of 15,451 words, representing 5.18% of the total corpus. The occurrences found after the analysis of the corpus are shown in Table 4, ordered by frequency. The first column shows the terms found in the corpus. The second column displays the occurrences found and the third, their percentages. Finally, the fourth column shows the categories to which the terms belong to.

Terms	Occurrences	(Percentage)	Categories
hotel	186	(23.2%)	facilities
habitación	155	(19.3%)	room
desayuno	54	(6.7%)	catering
personal	48	(5.9%)	staff
decoración	33	(3.0%)	facilities
limpieza	23	(2.8%)	facilities
con vistas	22	(2.7%)	facilities
recepción	21	(2.6%)	facilities
estancia	15	(1.8%)	management
negocios	13	(1.6%)	management
situación	13	(1.6%)	general
cama	13	(1.6%)	room
relación calidad-precio	13	(1.5%)	management
insonorización	12	(1.4%)	facilities
pega	11	(1.3%)	review platform
estrellas	10	(1.2%)	management
diseño	9	(1.1%)	facilities
instalación	9	(1.1%)	facilities
almohada	8	(0.9%)	room
aseo	7	(0.8%)	room
entrantes	7	(0.8%)	catering
reserva	7	(0.8%)	management
azotea	6	(0.7%)	facilities
comodidad	6	(0.7%)	facilities
ducha	6	(0.7%)	room
huésped	6	(0.7%)	management
recepcionista	6	(0.7%)	staff
barra	5	(0.6%)	facilities
lujo	5	(0.6%)	management
armario	5	(0.6%)	room

expectativas	4	(0.4%)	general
aire acondicionado	4	(0.4%)	facilities
casco antiguo	4	(0.4%)	leisure activities
colchón	4	(0.4%)	room
comentario	3	(0.3%)	review platform
empleados	3	(0.3%)	staff
fachada	3	(0.3%)	facilities
ingredientes	3	(0.3%)	catering
recomendaciones	3	(0.3%)	review platform
ubicación	3	(0.3%)	general
caja fuerte	3	(0.3%)	room
cuarto de baño	3	(0.3%)	room
escapada	3	(0.3%)	leisure activities
moqueta	3	(0.3%)	facilities
Tripadvisor	3	(0.3%)	review platform
referencias	2	(0.2%)	review platform
zona de la piscina	1	(0.1%)	facilities
ambientación	1	(0.1%)	facilities
artículos de tocador	1	(0.1%)	room
atracción turística	1	(0.1%)	leisure activities
cama supletoria	1	(0.1%)	room
confortabilidad	1	(0.1%)	room
enchufe	1	(0.1%)	room
lugares turísticos	1	(0.1%)	leisure activities
salones	1	(0.1%)	facilities
tarjetas de entrada	1	(0.1%)	technology
velocidad de internet	1	(0.1%)	technology
vestidor	1	(0.1%)	room
zona de baño	1	(0.1%)	facilities
zona de desayunos	1	(0.1%)	facilities
zona de estar	1	(0.1%)	facilities
zonas comunes	1	(0.1%)	facilities
Total	*801*	*(100.0%)*	

Table 4: Occurrences of the terms found in the corpus

At first sight, what seems immediately obvious is that the majority of consumer reviews are aimed at evaluating the hotel (186 occurrences) and the room (155 occurrences). The other three aspects that concern visitors the most are breakfast, hotel staff and décor. This simple frequency analysis reveals important information for our research on terminology in specialized texts, as the terminological density of online consumer reviews is lower than might be expected. Most of the vocabulary used in consumer reviews may be classified as unspecific of tourism.

We also considered it interesting to compare the frequencies obtained for each category. Figure 1 displays the occurrences of the specific terms classified in the nine categories, which are: facilities (e.g. *azotea, barra*), room (e.g. *almohada, armario*), management (e.g. *huésped, negocios*), catering (e.g. *desayuno, entrantes*), staff (e.g. *personal, empleados*), review platform (e.g. *referencias, recomendaciones*), general (e.g. *expectativas, situación*), leisure activities (e.g. *escapada, casco antiguo*) and technology (e.g. *tarjeta de entrada, internet*):
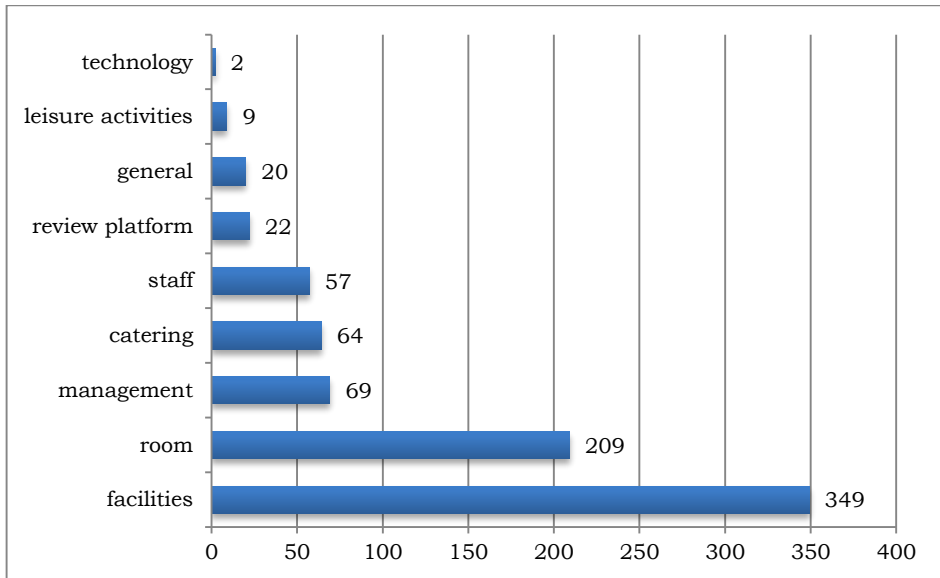
Figure 1: Comparison of the specific terms found in the corpus

Within the classification of the terms per specific category, shown in Figure 1, the facilities category contains the highest amount of occurrences in the corpus of consumer reviews, with 349 occurrences (44%), followed by that of room, with 209 occurrences (26%). Within the category of facilities, 232 terms were classified as technical and 117 as semi-technical, as can been seen in Table 4. However, this value is less important than it might seem, as hotel appears 186 times and it is usually at the beginning of the review: "*El hotel, está situado en...*", "*Muy buen hotel con una ubicación excelente...*", "*Me alojé en este hotel hace ya un tiempo...*", or "*Es la primera vez que me alojo en este hotel*".

In the next step of this study, we turned to the identification and classification of borrowing to study which ones were neologisms, as was explained in the methodology section of this paper. 22 borrowings were found in the corpus with a total number of occurrences of 34, out of 15,451 words, i.e. 0.22 % of the overall total. The results were lower than expected. The results are shown in Table 5, the first column shows the borrowings found, the second column displays the occurrences and the percentages calculated taking into account the total occurrences of borrowings (34) and the last column classifies them following the same classification in semantic categories used above to explain the terms found:

Borrowings	Occurrences (Percentage)	Semantic categories
Wi-Fi	6 (17.6%)	technology
web	2 (5.8%)	technology
check out	2 (5.8%)	accommodation

brunch	2 (5.8%)	catering
chill-out	2 (5.8%)	facilities
amenities	2 (5.8%)	room
mini bar	2 (5.8%)	room
suite	2 (5.8%)	room
premium	1 (2.9%)	room
mix	1 (2.9%)	facilities
check-in	1 (2.9%)	accommodation
sandwich	1 (2.9%)	catering
report	1 (2.9%)	review platform
staff	1 (2.9%)	staff
parking	1 (2.9%)	facilities
boutique	1 (2.9%)	facilities
cool	1 (2.9%)	facilities
fashion	1 (2.9%)	facilities
handicap	1 (2.9%)	facilities
standard	1 (2.9%)	room
deluxe	1 (2.9%)	room
jacuzzi	1 (2.9%)	room
Total	*34 (100.0%)*	

Table 5: Occurrences and semantic fields of the neologisms

The semantic domains of the borrowings found were also compared in order to gauge the importance of certain domains in the creation of new words. Only seven semantic fields were detected, we did not find borrowings that fit in the fields of management and leisure. The comparison of the occurrences found can be observed in Figure 2:
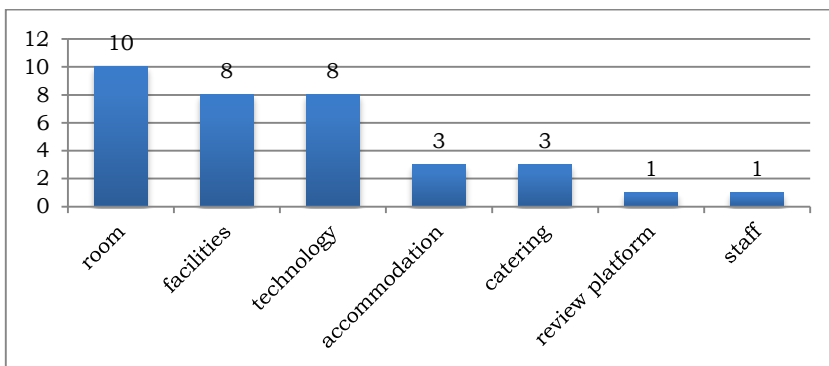

Figure 2: Classification of borrowings by semantic categories

The semantic field with the highest number of borrowings, 29% of the total, was that of room, followed by the semantics fields of technology (24%) and facilities (23%).

It should also be noticed the case of the borrowing Wi-Fi. According to the Oxford English Dictionary (2015), Wi-Fi is "(...) a proprietary name for: any of several standards for the high-speed wireless transmission of data over a relatively small range. Hence: a facility allowing computers,

smartphones, or other devices to connect to the Internet or communicate with one another wirelessly within a particular area by means of one of these standards". Although the spelling is consolidated as *Wi-Fi*, as it is a trademark of the Wi-Fi Alliance and a brand name for products using the IEEE 802.11 family of standards, in our corpus Wi-Fi varies in spelling or gender as observed in Example [4]:

"*El servicio de Wifi lamentable*", "*Hay wifi pero más vale armarse de paciencia*"; and change of gender *(la wifi, el wifi)*: "*Hecho en falta un wifi*", "*Con buena conexión Wifi*" or *"La wifi bastante lenta".*

The following step was the identification of neologisms from the borrowings shown in Table 5. After checking if the borrowings could be found in the *Diccionario de la Real Academia de la Lengua Española (DRAE),* the following terms used in consumer reviews may be considered neologisms (Table 6):

Neologisms	Occurrences	Semantic fields
check out	2	accommodation
brunch	2	catering
chill-out	2	facilities
amenities	2	room
premium	1	room
mix	1	facilities
check-in	1	accommodation
report	1	review platform
staff	1	staff
cool	1	facilities
fashion	1	facilities
deluxe	1	room

Table 6: Occurrences of neologisms found in the corpus

As can be observed, neologisms were not frequent in our corpus, but this result was expected given the domain under study and also the results presented by other researchers and mentioned above (Megerdoomian and Hadjarian, 2010, Cabré and Nazar, 2011).

Regarding the degree of specialization of the terms, the main explanation for the absence of specialized terminology in our corpus of consumer reviews could be that although this may be a specialized context, concerning tourism, the participants are not professionals or specialists, as Alcaraz (2000) pointed out with regard to such situations, concepts and knowledge are not then transmitted by specialized terminology, since the genre of consumer reviews is open to anyone who wishes to participate. For Calvi (2010: 23), in this informal genre, the tourist becomes the expert and creator of the discourse of tourism. After the analysis of the results, this seemed to be the main reason for the variation of terms found in the corpus. We found out that consumer reviewers preferred to use synonyms when expressing their opinions. Some examples can be seen in Table 7:

		Variation	in the words used	
Significados	1	2	3	4
Example [1]	Recepcionis-tas	Personal de re-cepción	Personal del hotel	Personas en los servicios de conserjería
Example [2]	Diseño	Decoración	Ambientación	
Example [3]	Cuarto de baño	Aseo	Baño	
Example [4]	Amenities	Artículos de toca-dor		
Example [5]	Salida	Checkout		
Example [6]	Opinión	Report	Comentario	Referencia

Table 7: Examples of variation in consumer reviews

We also noticed that the users of online consumer reviews tend to refer to the platform and the consumer review itself. Thus, visitors repeated the terms they read in the platform or in the questionnaire. For example, the use of the terms *referencia, comentario* or *recomendación*, in forty-one instances consumers mentioned their intention to recommend the hotel. The platform's invitation to the user "*You're helping travellers make better choices and plan their dream trips. Thank you!*" seems to be effective, as the reviewers used the same term that they had read in the platform. We believe that in this case, it cannot be established if reviewers used the terms as an individual action or if they were influenced by the terms used in the platform. Although our sample of consumer reviews is formed up of free-style texts, visitors closely follow the instructions given by the review platform and include a direct reference to their recommendation of the hotel as can be observed in example [1]:

"*Desde luego que es recomendable cien por cien*", "*Recomendable, pero muy caro...*", "*Muy recomendable*", "*Repetiría sin duda y lo recomendaría a todo el mundo*", "*Muy recomendable, tranquilo y bien ubicado.*"

Following the approach of Cabré (2006), all neologisms identified in this study fall into the category of lexical borrowings, with direct borrowings such as *amenities*, *brunch*, *suite*, or *premium*, or ones which undergo adaptation such as *check-out/check out*, appearing in the form of different spellings or as variants.

The term *staff* appears only once, but there are up to twelve different words or phrases used for the same concept in the corpus, probably due to the different social backgrounds of the users and their command of hotel terminology as can be seen in example [2]:

"(...) *personal de recepción, recepcionistas, personal, personal del hotel, chicos de recepción, personas en los servicios de conserjería, empleados* and *recepción del hotel*".

Both *check-in* and *check-out* are more commonly used as verbs in English, without the hyphen, but in Spanish they undergo a change of grammatical category to that of nouns, probably under the influence of their Spanish equivalents *hacer el registro (de entrada/salida)*. Examples are shown in [3]:

"(…) *muy amables tanto en el check-in como a la salida*" or with the Spanish verb *hacer* as in "(…) *a la hora de hacer check out (….) son muy lentos*".

However, in one of the reviews, the English term appears alongside an explanation in Spanish "*El hotel dispone de una interesante opción de 'late-check-out', es decir, la posibilidad dejar el hotel hasta las 17 horas*", but it is not clear whether the objective here is to explain the term *check-out*, or to specify the time range for a late check-out.

The dictionary of tourism and leisure terms (Alcaraz, Hughes, Campos, Pina and Aleson 2000) includes the record *amenity,* which is defined as service and comfort. In English, the plural form *amenities* is commonly used as the short form of *hotel amenities*, which in Spanish is usually translated as *productos de acogida* or *lotes de bienvenida*, referring to the toiletry products that hotels offer guests for added convenience. In the corpus, the manner of referring to this concept varies, as has been pointed out in Table 8, however there are no occurrences of the above-mentioned specialized terms but rather examples such as "(…) *cuenta con articulos de tocador de buenas marcas*", and the English term seems to prevail, although with changes in gender "(…) *muy buenos amenities*" and "(…) *las amenities perfectas*".

Another aspect to be mentioned, apart from neologisms, is that acronyms are scarcely used in our corpus, but one of the most common items did appear in relation to the issue of value for money, often referred to in Spanish with the phrase *relación calidad-precio*. This appears in 76 reviews, showing some inconsistencies and giving rise to the only acronym (RCP). Some examples that illustrate this variation are seen in [4]:
"*Relación calidad precio*", "*La relación calidad-precio fue bastante buena*", "*Calidad precio excelente*", "*Quizá la mejor RCP de Valencia*".

4. CONCLUSIONS

The conclusions drawn from the present study showed that within the classification of the terms per category of specialisation, shown in Figure 1, the category of *facilities* was the best represented in our corpus of consumer reviews with 44%, followed by *room*, with 26%. On the other hand, in the classification of the lexical borrowings, the most frequent category was that of *room* with 29% of the occurrences, followed by *technology* with 24%, and *facilities* with 23%.

It was observed that customers did not use highly specialised terms, possibly due to their social backgrounds and the medium of communication. Given the nature of tourism 2.0 consumer reviews, the frequency of neologisms found in the analysis does not seem to point to a trend in the creation of neologisms. It was also seen that user reviews are influenced by the guidelines and structure that the platform provides for writing the texts and in some cases from previous reviews by other users. As we found out in the corpus, some consumers repeat the same phrase:"*como otros han comentado/ (…) as someone said in an earlier comment*".

The comparison of the results for the lexical borrowings and the terms has revealed that sometimes tourism concepts are referred to in both ways, i.e. by lexical borrowing, with examples being: *amenities/productos de tocador, check out/salida, staff/personal*. In other cases, the concept is referred to only by means of a lexical borrowing, as in cases such as *mini bar, Wi-Fi, brunch, jacuzzi, chill-out,* or *boutique*, for example. This might entail that Spanish tourists use some English words in consumer reviews that have been incorporated from the vocabulary used in hotels or in journeys.

Unquestionably, the data obtained from online reviews and textual consumer-generated content represent a reliable measurement of customer satisfaction, and therefore further research on the analytical tools needed to process and transform these data would be beneficial to both customers and hoteliers. We are conscious that the data extracted for this analysis should be expanded as more consumer reviews might give us a wider scope of the use of neologisms and English borrowings in Spanish.

REFERENCES

ALCARAZ, E. (2000): *El inglés profesional y académico*, Madrid: Alianza Editorial.

ALCARAZ, E., HUGHES, B. CAMPOS, M. A., PINA, V. & ALESÓN, A. (2000): *Diccionario de términos de turismo y ocio. Inglés-español, Spanish-English,* Barcelona: Ariel.

BARBU, A. M. (2011): "Monitoring neologisms in newspapers within the NEOROM project", *Memoirs of the Scientific Sections of the Romanian Academy. Tome XXXIV*, pp. 1-17.

CABRÉ, M. T. (1999): *Terminology. Theory, methods and applications,* Amsterdam: John Benjamins Publishing Company.

CABRÉ, M. T. (1998): "Las fuentes terminológicas para la traducción", Fernández Nistal, P. and Bravo, G. J. (eds.): *La traducción: Orientaciones lingüísticas y culturales,* Valladolid: Universidad de Valladolid, pp. 27-59.

CABRÉ, M. T. (2002): *La terminología. Representación y comunicación*, Barcelona: UPF.

CABRÉ, M. T. (2003): "Theories of terminology: Their description, prescription and explanation", *Terminology,* 9(2), pp. 163-199.

CABRÉ, M. T. (2006): "La clasificación de neologismos: una tarea compleja", *Alf*a, 50(2), pp. 229-250.

CABRÉ, M. T. & ESTOPÀ, R. (2004): *Metodología del trabajo en neología: criterios, materiales y procesos*, Barcelona: Institut Pompeu Fabra.

CABRÉ, M. T. & NAZAR, R. (2011): "Towards a new approach to the study of neology", *4th Joint Seminar on Neology and specialised translation*, Brussels: CVC and Termisti.

CABRÉ, M. T., ESTOPÀ, R. & VARGAS, Ch. (2012): "Neology in specialized communication", Special issue of *Terminology,* 18(1), pp. 1-8.

CALVI, M. V. (2001): *El léxico del turismo. Culturele*, Barcelona: Universitat de Barcelona. [Accessed on 25/03/2016: http://www.ub.edu/filhis/culturele/turismo .html].

CALVI, M. V. (2010): "Los géneros discursivos en la lengua del turismo: una propuesta de clasificación", *Ibérica,* 19, pp. 9-32.

CARRIÓ-PASTOR, M. L. & CANDEL MORA, M. A. (2013): "Variation in the translation patterns of English complex noun phrases into Spanish in a specific domain", *Languages in Contrast,* 13(1), pp. 28–45.

CARRIÓ-PASTOR, M. L. (2015): "Identification of rhetorical moves in business emails written by Indian speakers of English", Darics, E. (ed.), *Digital Business Discourse,* London: Palgrave Macmillan Publishers, pp. 226-242.

COX, C. *et al.* (2008): *Consumer-generated web-based tourism marketing,* Queensland: CRC for Sustainable Tourism Ltd.

DARICS, E. (ed.) (2015): *Digital Business Discourse*, London: Palgrave Macmillan Publishers.

DUBUC, R. & LAURISTON, A. (1997): "Terms and Contexts", Wright, S. E. and Budin, G. (Eds.), *Handbook of Terminology Management,* Amsterdam: John Benjamins Publishing, pp. 80-88.

EDO, N. (2012): "Páginas web privadas e institucionales: el uso de la adjetivación en un corpus inglés-español de promoción de destinos turísticos", J. Sanmartín (ed.), *Discurso turístico e Internet,* Madrid: Interamericana, pp. 15-25.

GAINS, J. (1999): "Electronic mail - A new style of communication or just a new medium?: An investigation into the text features of e-mail", *English for Specific Purposes,* 18(1), pp. 81-101.

GOETHALS, P. (2014): "La acomodación lingüística en contextos profesionales turísticos. Un enfoque didáctico basado en los testimonios de turistas", *Ibérica,* 28, pp. 181-202.

GOTTI, M. (2003): *Specialized Discourse. Linguistic Features and Changing Conventions*, Bern: Peter Lang.

KESSLER, G. (2010): "Virtual business: An Enron email corpus study", *Journal of Pragmatics,* 42, pp. 262-270.

LERAT, P. (1997): *Las lenguas especializadas*, Barcelona: Ariel.

MEGERDOOMIAN, K. & HADJARIAN, A. (2010): Mining and classification of neologisms in Persian blogs. *Proceedings of the NAACL HLT 2010*: 6-13.

MOGHADAM, M. Y. & SEDIGHI, A. (2012): "A study of the translation of neologisms in technical texts: a case of computer texts", *International Journal of Scientific & Engineering Research,* 3(2), pp. 1-6.

Oxford English Dictionary (1999): [Accessed on 23/03/2016: http://www.oed.com].

O'REILLY, T. (2007): "What is web 2.0: Design patterns and business models for the next generation of software", *Communication and strategies,* 65, pp. 17-37.

RESCHE, C. (2000): "An Approach to Interface Terminology: The Example of Environmental Economics in English as a Foreign Language", *Meta* XLV (4), pp. 628-645.

SAGER, J. C. (1990): *A practical course in terminology processing*, Amsterdam: John Benjamins Publishing Company.

SAGER, J. C. (1997): "Term Formation", Wright, S. E. and Budin, G. (Eds.), *Handbook of Terminology Management, Volume I. Basic Aspects of Ter-*

minology Management, Amsterdam: John Benjamins Publishing, pp. 27-41.

SCHEMMANN, B. (2011): "A Classification of Presentation Forms of Travel and Tourism-Related Online Consumer Reviews", *e-Review of Tourism Research,* 9, pp. 14-24.

SCOLLON, R. & SCOLLON, S. W. (2004): *Nexus analysis: Discourse and the emerging internet*, London: Routledge.

SEARGEANT, P. & TAGG, C. (eds.) (2014): *The Language of Social Media: Identity and Community on the Internet*, London: Palgrave MacMillan Publishers.

SECRETARIAT OF STATE FOR TOURISM OF FRANCE. (2001): *Thesaurus on Tourism & Leisure Activities*, Paris: World Tourism Organization.

SKOVHOLT, K. & SVENNEVIG, J. (2013): "Responses and non-responses in e-mail interaction", Herring, S., Stein, D. and Virtanen, T. (eds.), *The*

Pragmatics of Computer-Mediated Communication, Berlin: De Gruyter Mouton, pp. 589-612.

TALEBINEJAD, M. R., DASTJERDI, H. V. & MAHMOODI, R. (2012): "Barriers to technical terms in translation", *Terminology*, 18(2), pp. 167-187.

TEMMERMAN, R. (2000): *Towards new ways of terminology description. The sociocognitive approach*, Amsterdam: John Benjamins Publishing Company.

VÁSQUEZ, C. (2014): *Online consumer reviews*, New York: Bloomsbury.

WORSOE, L. B. (2011): "What's in a word - what's a word in?", Language *Sciences,* 33, pp. 603-613.

WÜSTER, E. (1974): "General Terminology Theory - Fine Line between Linguistics, Logic, Ontology, Information Science and Business Sciences", *Linguistics,* 119, pp. 61-106.