JoSTrans
The Journal of Specialised Translation

Combinatorial patterns in medical case reports: an English-Spanish contrastive analysis[1].

Beatriz Méndez-Cendón, ITBYTE, Universidad de Valladolid, Spain.

ABSTRACT

There have been several studies related to the description of phraseological and rhetorical structures in specific types of medical prose genres. However, to our knowledge, these studies have never focused on combinatorial patterns and have never been interlinguistic. In this paper we analyse the phraseological patterns of medical case reports published in journals and online forums. It is our purpose to identify Spanish and English combinatorial patterns which are recurrent in the discourse of this text genre so as to establish intralinguistic and interlinguistic equivalences among them. We also account for the role phraseological patterns play in the building of information in case reports and in the communicative aim of this genre. We firmly believe that knowing the phraseological practices of the genre and its internal ordering of information is essential to produce coherent and readable target texts.

KEYWORDS

LSP phraseology, combinatorial pattern, LSP genre, case report, electronic corpus.

1. Introduction

This paper proposes a corpus-based analysis to retrieve the phraseology specific to medical case reports. Our proposal emphasises the existence of a set of conventional phrasal practices in the writings of this text genre. The underlying assumption of our study is "the idiom principle" (Sinclair, 1991: 110), that is, words do not occur at random in texts but they show a tendency to co-occur with certain grammatical choices and in certain semantic environments.

We will adopt a genre-analysis approach in the description of the phraseology (following Swales, 1990 and 2004) and we start out with the hypothesis that the language used in very formal medical texts contains a large number of partly prefabricated expressions which are used according to the communicative purposes of the genre.

Our purpose is to examine the textual role these fixed expressions play in the chosen genre and whether they are conditioned by its rhetorical aims. We propose to analyse their linguistic and textual behavior in order to develop some methodological guidelines that could be of use to translators and could also be applied to translation teaching materials for this type of text.

In the following section we discuss the concepts of 'language for specific purposes (LSP) phraseology' and 'text genre' since these are key concepts to situate the results of the phraseological analysis we perform in this paper.

1.1. LSP phraseology

Generally speaking, LSP phraseology is the study of specialised word combinations that occur frequently in technical and scientific language. These combinations are termed in different ways by different authors. For example, 'phrasal lexeme', 'set expression', 'multiword unit', 'phraseme', 'LSP phrase' and 'phraseological unit' have all been used.

LSP phraseological units (PUs) can be defined as recurrent word combinations which occur in specialised language. PUs are characterised by a high frequency of cooccurrence of their constituent elements and the semantic and syntactic connections established between them. These constituent elements are fixed but the PU admits internal variation, for example, the permutation of the elements, or the substitution of one element for another. Examples of specialised PUs in medical language are collocations, irreversible binomials, routine formulae and combinatorial patterns.

There are different types of lexical units which give rise to PUs in medical language:

(1) Terms which are specific to a medical field, or a medical specialty, for example 'angiography', 'tomography' and 'radiograph' belong to the field of diagnostic imaging. They form collocations such as 'to perform radiographs' or 'to undergo angiography'.
(2) Interdisciplinary terms, that is, terms belonging to other related fields such as physics, for example 'echo', 'magnetic resonance' and 'pulse sequence'. They are part of combinations occurring in the language of diagnostic imaging, such as 'to acquire sequences' or 'spin-echo imaging'.
(3) Subtechnical, or semi-technical terms, coming from general language which are used in medical language with a more limited and precise meaning, such as 'examination', 'case', 'evidence', 'test' or 'findings'. Collocations such as 'to examine a patient' or 'to show evidence of' are very frequent in medical writings.
(4) Finally, general language words such as 'man', 'woman', 'time' or 'number'. These words can be found as constituents of grammatical collocations such as 'at the time of' or 'a 75-year-old-woman/man with a history of'.

All these lexical units are integrated at the syntactic level and make up the phraseology of a medical field.

1.2. Medical case reports

Medical case reports are one of the most popular text genres in medical academic prose[2]. They describe some new or rare finding associated with a disease, syndrome or disorder found in a patient. Journals may differ

widely in their policies regarding the case report layout (superstructure), but in general the format is: 'Introduction', 'Case Description', and 'Discussion'.

The Introduction states why the case is unique or unusual and how it contributes to medical knowledge. The Case Description sums up information found in the patient's files, including present state of health and past medical history, findings from physical examinations and laboratory tests, differential diagnosis, treatment and follow-up. The Discussion evaluates the patient's symptoms, provides evidence to support the diagnosis and suggests implications for clinical practice and possibilities of further research.

Medical case reports are mainly published in two media: medical journals and specialty Internet forums. Medical journals have each their own guidelines regarding the publication of case reports. In general, these printed case reports share a particular layout:

(1) Abstract
(2) Introduction
(3) Case report
(4) Discussion
(5) References.

They also have similar communicative aims, namely to convince the reader of the existence of an unusual syndrome or disease, to provide supporting evidence, to suggest an adequate treatment, and to make recommendations for further study.

As regards case reports submitted to online forums, they usually have a different layout:

(1) Patient History
(2) Gross and microscopic description
(3) Final Diagnosis
(4) Discussion.

Their purpose is to present other medical specialists from around the world with an uncommon disease, syndrome or disorder and ask for a final diagnosis. Both layouts are different because they have different communicative purposes. Case reports from journals follow the conventional academic order of the written research paper since the purpose is the dissemination of research within the medical community, whereas case reports submitted to forums are written by different types of medical specialists who are seeking a diagnosis for a clinical case found in their daily practice. This search gives rise to discussion in the medical community.

Both types of publication, however, have in common the section presenting the patient's case. This case description is called Case Report section in journals and Patient History section in electronic forums (from now on referred to just as the Case Report section).

A preliminary research conducted by the author (Méndez-Cendón, 2005) showed that authors of case reports use a formulaic language to present the patient's case in printed and online case reports. Therefore, we will be focusing on the phraseological analysis of the Case Report section in this paper.

1.2.1. The Case Report section

Our preliminary work also revealed that the organisation of information in this section is always structured into four different moves according to Nwogu's terminology (1997)[3].

1. Presenting the patient's symptoms.
2. Describing physical examination findings.
3. Describing laboratory and other test findings.
4. Describing diagnosis, treatment and follow-up.

The first move presents information about the patient in chronological order: the patient's signs and symptoms, chief complaint, medical history, any relevant family history, social history and medication being taken. The second move reveals the results of physical examinations. The third move shows laboratory tests and diagnostic imaging findings. Furthermore, this third move includes visual elements that mainly explain imaging findings. Finally, move four explains the final diagnosis, the treatment and outcome of treatment. The identification of these four moves allowed us to classify the combinatorial patterns detected in our study as belonging to each one of them.

2. Research hypothesis and objectives

The working hypothesis of this study was that authors of case reports might use prototypical phraseological features to introduce relevant information in the Case Report section since the use of a formulaic language is its predominant linguistic feature (cf. Méndez Cendón 2005). More specifically, authors would use prototypical phraseological features in case reports written in English and Spanish, and published both in paper format and electronically.

The objectives to be achieved in this paper in order that the working hypothesis could be corroborated were:

1)   To carry out a phraseological analysis of the Case Report section and identify different combinatorial patterns starting from certain lexical units which are predominant in this section.

2)   To show similarities or differences between the patterns used in English and Spanish journals and online forums, since we believe that the type of language, medium of diffusion and communicative purposes of the case report may impose some linguistic restrictions.

3)   To offer specialised translators some phraseological guidelines specific to the genre, the language and the type of publication.

3.   Methodology

3.1.   Compiling the Case Report Corpus (CRC)

The first step in our research was compiling the Case Report Corpus (from now on referred to as CRC). In selecting case reports we took into account their availability in electronic format, as well as whether they were published in popular medical journals. The reports compiled in our corpus were extracted from printed medical journals (from now on referred to as PCRs) and electronic medical forums (ECRs).

The PCRs were obtained from two English medical journals: *Radiology* (the Official Publication of the Radiological Society of North America) and *European Journal of Internal Medicine* (The Official Publication of the European Federation of Internal Medicine). They were also selected from two Spanish journals: *Radiología* (the Official Publication of the Spanish Radiological Society) and *Medicina General* (the Official Journal of the Spanish Society of Family Doctors). We chose these journals because they have a similar audience in English and Spanish medical communities and were available online. While *Radiology* and *Radiología* target experts in radiology, *European Journal of Internal Medicine* and *Medicina General* are essential readings for all type of medical practitioners, especially generalists and family doctors, specialists in internal medicine and clinical researchers. The fact that there are two different types of readership for these journals could provide us with interesting intralinguistic comparisons between the texts.

Regarding the ECRs, the English ones were published in two medical forums online: *The University of Pittsburgh Online Case Studies* and *Medscape Clinical Cases*. As for the two Spanish forums selected online, these were *Eu-Salud* and *Foro de Diagnóstico por Imagen*. In this way we were able to obtain a great number and variety of instances. Therefore, the CRC comprises two bilingual comparable subcorpora: the PCR corpus and the ECR corpus. The composition of these two subcorpora is shown in Tables 1 and 2 by name of publication, Internet site, language, year and number of words.

Name of publication	Language	Year	Number of words in the *Case Report* section
European Journal of Internal Medicine (http://www.elsevier.com)	ENGLISH	2004	5,604
Radiology http://radiology.rsnajnls.org/	ENGLISH	2004	6,202
Medicina General (http://www.medicinagene ral.org/)	SPANISH	2004	7,347
Radiología (http://db.doyma.es/cgi-bin/wdbcgi.exe/doyma/mre vista.salta_a_ultimo?pident =119)	SPANISH	2004	6,010
TOTAL NUMBER OF WORDS			25,163

Table 1: Composition of the PCR corpus.

Name of publication	Language	Year	Number of words in the *Patient History* section
University of Pittsburgh Online Case Studies (http://path.upmc.edu/ca semonth.html)	ENGLISH	2004	6,035
Medscape Clinical Cases (http://www.medscape.c om/medgenmed-home)	ENGLISH	2004	5,730
Eu-Salud (http://eusalud.uninet.ed u/)	SPANISH	2004	5,010
Foro de discusión Elmédico.net (http://elmedico.metropo liglobal.com/)	SPANISH	2004	5,130
TOTAL NUMBER OF WORDS			21,905

Table 2: Composition of the ECR corpus.

In summary, both PCR and ECR subcorpora were compiled following identical selection criteria so as to obtain relevant phraseological data in English and Spanish languages.

3.2.   Identifying the phraseology

The next step was to extract the phraseology from our corpus using WordSmith Tools v. 4.0. The WordList Tool option allowed us to extract

the words most frequently used in the Case Report section (see Figure 1 below).

N	Word	Freq.	%	N	Word	Freq.	%
1	PATIENT	316	0,33	2	CASE	285	0,30
3	**FIG**	264	0,28	4	SHOWED	204	0,21
5	RIGHT	190	0,20	6	HISTORY	177	0,19
7	LEFT	170	0,18	8	CELLS	165	0,17
9	PATIENTS	165	0,17	10	CALCIFICATIONS	160	0,17
11	**EXAMINATION**	160	0,17	12	MASS	160	0,17
13	NORMAL	153	0,16	14	**FIGURE**	151	0,16
15	**YEAR**	148	0,16	16	MG	145	0,15
17	SHOWS	145	0,15	18	**AÑOS**	139	0,15
79	CHEST	60	0,06	80	MAGNIFICATION	60	0,06
81	**FIGURA**	59	0,06	82	HOWEVER	59	0,06
123	OBSTRUCTION	51	0,05	124	HACE	50	0,05
125	MARKED	50	0,05	126	SEEN	50	0,05
127	ANTECEDENTES	49	0,05	128	IT	49	0,05
129	MATERIAL	49	0,05	130	MUJER	49	0,05
131	ESTUDIO	48	0,05	132	IMAGEN	48	0,05
137	**EXPLORACIÓN**	47	0,05	138	GROSS	47	0,05
139	LOW	47	0,05	140	STAIN	47	0,05
141	**EVOLUCIÓN**	47	0,05	142	UP	47	0,05
143	HAS	46	0,05	144	MORE	46	0,05
145	PATHOLOGY	46	0,05	146	RISK	46	0,05
147	WOMAN	46	0,05	148	ARTERIES	45	0,05
149	CASES	45	0,05	150	IMAGES	45	0,05
151	SPECIMEN	45	0,05	152	VASCULAR	45	0,05
153	DAY	44	0,05	154	**FOLLOW**	44	0,05

Figure 1. The most frequent words used in the *Patient History / Case Report* section.

At this point, we selected certain lexical units and assigned them to each one of the four moves of this section. The criteria used in selecting these units were their meanings and their high frequency of occurrence in the case report section. This way we assigned 'year' and *años* to the first move; 'examination' and *exploración* to the second move; 'figure'*, figura* and the abbreviated form 'fig'*.* to the third move and, finally, 'follow-up' and *evolución* to the fourth move.

Using the Concord Tool option we were able to get KWIC concordance lines of the lexical units previously selected. A syntactic analysis of these lines showed different fixed and conventionalised expressions presenting a similar syntactic behavior. A further semantic analysis revealed that these fixed expressions could be grouped according to their meaning into different combinatorial patterns[4]. The analysis of these patterns required grouping sets of terms and meaning-related expressions into the following six *conceptual areas* or *categories*, all of them closely interrelated:

(1) SYMPTOM. This category comprises terms describing patients' symptoms, such as 'feve', 'vomitus' and 'pain'*.*

(2) SIGN. Terms referring to patients' signs are included in this category. For example, 'splenomegaly', 'cyanosis', and 'jaundice'.

(3) DISEASE. This category includes terms designating an illness, such as 'tumor' and 'neoplasm'.

(4) FINDING. This category refers to terms describing abnormalities found on diagnostic images, such as 'hypoechoic masses', 'hypoattenuating structures' and 'hypervascularised hepatic lesion'.

(5) DIAGNOSTIC IMAGE. This category consists of terms denoting types of diagnostic images, such as 'CT scan', 'MR image' and 'radiograph'.

(6) DIAGNOSTIC EXAMINATION. Finally, terms referring to types of diagnostic studies performed on patients belong to this conceptual area. For example, 'MR imaging', 'MR angiography' and 'conventional hysterosalpingography'.

Finally, our phraseological analysis revealed the use of several combinatorial patterns specific to the first three moves of the Case Report section. Thus, the results of the analysis presented in section 4 are focused only on those three moves.

4. Results

4.1. Presenting the patient's symptoms

There are four lexical items specific to this first move: 'man'*, varón* (Spanish for 'man'), 'woman' and *mujer* (Spanish for 'woman')*.* They occur repeatedly in both PCRs and ECRs to introduce different combinatorial patterns according to the type of publication.

The PCR corpus

Two typical combinatorial patterns were found in Spanish journals at the beginning of this move:

(i) *Varón/mujer de … años de edad con antecedente(s) de/diagnosticado(a) de* DISEASE *que presenta/acude a/consulta por* SYMPTOM.

(ii) *Se trata de un/una varón/mujer/paciente de … años de edad que presenta/acude a/ consulta por* SYMPTOM.

The occurrence of the first pattern is very high in Spanish PCRs (almost 84.7% of the total number of occurrences of both patterns). The following three examples illustrate the use of different Spanish verbs in the relative clause included in the pattern.

(1) Varón de 72 años de edad con antecedente de carcinoma de colon intervenido hace dos años <u>que presenta</u> bultoma en el brazo derecho.

(2) Varón de 56 años de edad, diagnosticado hace 8 años de HTA <u>que acude a</u> consulta <u>por</u> haber sufrido 15 días antes, de madrugada, un accidente cerebrovascular con un cuadro de hemiparesia derecha y afasia del que ya no se apreciaban secuelas.
(3) Mujer de 74 años de edad con antecedente de histerectomía y anexectomía <u>que consulta por</u> dolor abdominal de predominio en fosa ilíaca izquierda (FII).

The second pattern does not include the patient's medical history and its frequency is relatively low (15.2%), for example:

(4) Se trata de una mujer de 70 años de edad que consulta por infecciones urinarias de repetición. En la ecografía abdominal se descubre una gran masa heterogénea.

As for the English publications, the following combinatorial patterns were observed to introduce the patient's symptoms:

(i) *A …–year-old man/woman/male/female patient was admitted (to hospital/ pediatric service/ intensive care unit, etc.) with SYMPTOM.*
(ii) *A ....-year-old man/woman/male/female patient presented (at/to outpatient clinic/ hospital, etc.) with* SYMPTOM.

The first pattern occurs more frequently than the second one (52% of total occurrences) and includes a passive structure to introduce the patient's condition. More specifically, the grammatical collocation 'admit with' is very typical of English case reports, e.g.:

(5) A 63-year-old Caucasian woman <u>was admitted to</u> our hospital <u>with</u> a 2-month history of intense fatigue, weight loss (4 kg) and bone pain, most prominent in the knees.

The second pattern contains the grammatical collocation 'present with', which is a typical collocation in English medical language to introduce the patient's disorder, for example:

(6) A 77-year-old man <u>presented with</u> cardiac failure but failed to respond to conventional anti-failure treatment.

The place of admission can be specified optionally by the prepositions *at* or *to*:

(7) A 23-year-old female patient <u>presented at</u> the outpatient clinic <u>with</u> chronic fatigue.

The ECR corpus

The combinatorial patterns in Spanish ECRs also occur at the beginning of the Historia Clínica' section:

(i) *Varón/mujer/paciente de ... años con* SYMPTOM. For example:

(8) Mujer de 53 años con lesión nodular en partes blandas de borde cubital de mano derecha.

(ii) *Varón/mujer/paciente de ... años con/sin antecedentes de* DISEASE *que consulta por* SYMPTOM, e.g.:

(9) Paciente mujer de 68 años de edad, con antecedentes de hernia de hiato y ulcus péptico, que consulta por presentar en la parte inferior de la espalda desde hace 4 años unas líneas horizontales palpables.

The first combinatorial pattern occurs repeatedly in Spanish ECRs (70%). Both patterns are more concise than the patterns from Spanish PCRs.
As to the English ECRs, there are three combinatorial patterns identified in the corpus:

(i)   *The patient is/was a …-year-old man/woman who presents/presented with* SYMPTOM, for example:

(10) The patient is a 66-year-old man who presented with a left thyroid mass.

(ii) *This …-year-old man/woman presents/presented with* SYMPTOM, e.g.:

(11) This 53-year-old woman presented with a progressively enlarging, painless mass in the area of the left parotid gland.

(iii) *This was a ...-year-old man/woman who presented* (*to outpatient clinic/hospital, etc*) *with* SYMPTOM, for example:

(12) This was a 67-year old man who presented to the emergency department with worsening dyspnea and cough for 8 months and new onset chest pain.

The first pattern is the most popular in ECRs since it accounts for 80% of the pattern occurrences. In addition, the data shows that the simple past is preferred to the simple present in the first two patterns.
Sometimes the words 'male' and 'female' are used interchangeably with 'man' and 'woman' in English ECRs, e.g.:

(13) The patient was a 44 year-old male who presented with a chief complaint of a left flank mass which was gradually increasing in size.

It is worthwhile pointing out the use of the indefinite and definite articles in both PCRs and ECRs since it reveals a greater or lesser degree of specificity in language. In cases published in journals the indefinite article *a* denoting a generic reference is preferred for presenting the case description (see examples 5, 6, 7 above). On the other hand, the definite article (the' and the demonstrative 'this' convey a specific reference to the noun phrase identifying the patient (see examples 10, 11, 12, 13).

Finally, there is a substantial difference between verb tenses in the Spanish and English patterns introducing this first move. Whereas Spanish prefers the simple present to introduce the patient's symptoms, English uses the simple past (the occurrence of the verbs *acudir* and *consultar* is

almost 85% of the total occurrences in the simple present, whereas the occurrence of *present* in the English PCR and ECR corpora is very high in the past tense; being 95% of all occurrences of this verb).

4.2. Describing physical examination findings

The key terms specific to this second move are *examination* and its Spanish equivalent term *exploración* (see figure 1 above). These two terms collocate with nouns or adjectives to form specific noun groups. We will be focusing on the combinatorial patterns including the compound terms 'physical examination' and *exploración física* since they are extremely frequent in the corpus.

The PCR corpus

There are two typical combinatorial patterns in the Spanish journals to present the findings from medical tests:

(i) *En/a la exploración física el/la paciente presenta(ba)/se aprecia(ba)/se detecta(ba)/destaca(ba)* SIGN, e.g.:

(14) A la exploración física la paciente presentaba un aceptable estado general con una temperatura axilar de 38°C.

(ii) *La exploración física muestra/mostraba/era/revela/revelaba* SIGN, for example:

(15) La exploración física muestra un abdomen globuloso, con semiología de ascitis.

The first pattern is predominant (69.2%) and includes the prepositional phrases *en la exploración física* and *a la exploración física*. This last phrase is considered a specific grammatical collocation in the discourse of Spanish case reports to describe physical examination findings. In addition, the subject *paciente* can be replaced by a 'passive reflexive' clause as shown in the following example:

(16) A la exploración física <u>no se detectan</u> alteraciones del balance articular de la columna vertebral en de las extremidades.

Passive reflexives (*pasiva refleja)*, i.e. reflexive verbs used with a passive meaning, are very frequent in Spanish scientific prose. They make up a very productive special type of transitive clause used for depersonalisation, since the agent of the action is not explicitly stated. They are signaled by the pronoun *se*. There may be arbitrary switches between the simple present and the imperfect (*pretérito imperfecto)* verb tenses in Spanish printed case reports to describe the patient's examination. For example:

(17) HISTORIA CLÍNICA. Varón de 65 años de edad, con antecedentes personales de diabetes mellitus que <u>precisa</u> insulina para su tratamiento. <u>Consulta</u> por tumefacción y deformidad de los dedos de mano izquierda. A la exploración se <u>apreciaba</u> tumefacción de 2º y 3º dedos con lesiones de consistencia semidura indoloras a nivel de interfalángica distal y media. En analítica <u>destaca</u> glucemia: 259 mg/dl y Hb A1C 8,9.

Two typical patterns are used to introduce test findings in the English journals:

(i)   *Physical examination revealed/showed/disclosed* SIGN, for example:

(18) Physical examination revealed dullness at percussion and decreased air entry in the base of the left lung.

(ii)   *At physical examination* SIGN *was detected/was seen*:

(19) At physical examination, a palpable mass was detected in the patient's left upper quadrant.

The first pattern, which is the most usual one (almost 78%), includes an inanimate subject with an active verb (reveal, show, disclose). Thus, the term 'physical examination' seems to play the role of an active agent in the clause.

The second pattern includes a phrase that comprises the grammatical collocation 'at physical examination', which is very characteristic of the language of the journal *Radiology* when introducing findings from examinations.

The ECR corpus

Two combinatorial patterns are present in the discourse of Spanish ECRs:

(i)   *En la exploración física se percibe* SIGN, e.g:

(20) En la exploración física se percibe una ligera hepatoesplenomegalia. Existía el antecedente de estancias prolongadas y frecuentes.

(ii)   *Exploración física:* SIGN, e.g.:

(21) Exploración física: TA:200/110 mmHg, consciente, orientado, apirético y sin adenopatías periféricas.

The first pattern is more popular than the second one (80%). Again the linguistic structure is more concise than in the PCRs. In English ECRs, there is just one combinatorial pattern to explain findings from physical examinations:

Physical examination revealed/showed SIGN, for example:

(22) Physical examination also showed an enlarged axillary lymph node.

Finally, the simple present is the verb tense of choice in both Spanish PCRs and ECRs to introduce this second move, whereas English PCRs and ECRs are always in the past for this move.

4.3. Describing laboratory and other test findings

This third move is used to introduce laboratory and imaging test findings. It includes visual (non-verbal) elements to introduce the original findings obtained from the different diagnostic procedures such as x-rays, CT scanning or MR imaging. Visual elements in case reports provide data to convince readers of the validity of the findings and allow them to see how the data was obtained. Authors use two mechanisms to make reference to these non-verbal elements:

A.   References in the text to images. Visual elements are located in two positions, according to Busch-Lauer (1998: 116):
•   'Pre-text position': the visual elements occur before they are referred to in the text.
•   'Parallel position': the textual reference and the images are on the same page very close to each other.
B.   Captions included in the images.

We will be following this former division in the analysis of our data.
The terms 'figure' in English and its equivalent *figura* in Spanish are the lexical units which explicitly signal the information contained in this move (see section 3.2). 'Figure' and *figura* are written in two different ways, according to the type of information introduced. If the text refers to images, they are abbreviated to 'fig' and written in parentheses; however, if they are introducing a caption for an illustration, they are either abbreviated to 'Fig.' in both languages or written out in full. These abbreviations are usually followed by a numeral, e.g. '*Fig. 7*'.

The PCR corpus

A.   The text refers to images

Two combinatorial patterns are used to describe the disorders shown on images:

(i) *En el/la* DIAGNOSTIC IMAGE (*Fig. number*) *se evidencia/se observa/se objetiva/se visualiza/ se aprecia* FINDING, for example:

(23) En la TC de tórax con plano de corte supraaórtico y subcarinal (Figs. 2A y B) se observa asimetría de partes blandas a nivel supraaórtico por agenesia del músculo pectoral menor izquierdo.

(ii) *El/la* DIAGNOSTIC IMAGE *muestra/revela/evidencia* FINDING *(Fig. number)*, for example:

(24) La TC realizada a las 24 horas muestra un área hipodensa de morfología triangular, con cierto grado de expansividad en la región occipital derecha (Fig. 3).

The first pattern is the one preferred by Spanish doctors (73%) in their papers and includes a passive reflexive clause. The Spanish verbs *evidenciar*, *observar*, *objetivar*, *visualizar* and *apreciar* have the same meaning, namely that the lesion is exposed to view. In the corpus, *observar* is the most frequent verb, followed by *visualizar* and *apreciar*. In addition, both patterns include a parenthesis with the figure number, which is located in a different position in the sentence according to the pattern. On the other hand in English PCRs there is only one combinatorial pattern:

DIAGNOSTIC IMAGE showed/revealed/demonstrated/confirmed FINDING (Fig. number), e.g.

(25) Skeletal radiographs revealed severe diffuse osteoporosis, with cancellous bone hyperdensity and cortical bone hypodensity (Fig. 1).

The verbs 'show', 'reveal', 'demonstrate' and 'confirm' are used interchangeably as active verbs in this move in order to avoid repetition of the same word. With this last verb the authors affirm the validity of the information given as in the following example:

(26) A CT scan <u>confirmed</u> the presence of lymphadenopathies in the mediastinum and the hilar regions (Fig. 1).

B.   The caption is included in the images

The style guides for *Radiology, Radiología, European Journal of Internal Medicine* and *Medicina General* follow *The Uniform Requirements for Manuscripts Submitted to Biomedical Journals*, which are available at http://www.icmje.org/. These requirements state that images should be identified and explained clearly in the captions for illustrations, so that the reader will be able to understand them. In other words, the information represented in the figures should be translated into an explanation in the accompanying text.

In Spanish PCRs the captions exhibit the following patterns, listed by frequency of occurrence:

(i)   DIAGNOSTIC IMAGE. FINDING, for example:

(27) Fig. 3. Imagen sagital potenciada en T1 con técnica de saturación grasa. Lesión hemorrágica en la pantorrilla izquierda que presenta una pseudocápsula hipointensa a su alrededor.

(ii) DIAGNOSTIC IMAGE *que muestra/presenta* FINDING, for example:

(28) Fig.1.–Ecografía abdominal que muestra masa de predominio quístico con áreas sólidas esféricas.

(iii) DIAGNOSTIC IMAGE *en el/la que se demuestra/se visualiza/se aprecia/se observa* FINDING, for example:

(29) Fig.5.–TC abdominal con contraste IV a nivel del lóbulo hepático en la que se demuestra que la masa está en contacto con la pared gástrica.

Our research shows that the first pattern is the preferred one in both Spanish publications (68.5%).
As for the English PCRs, they include the following combinatorial patterns in their captions:

(i) DIAGNOSTIC IMAGE *obtained at/in* DIAGNOSTIC EXAMINATION *shows /reveals/demonstrates* FINDING, for example:

(30) Fig. 5b. Radiograph obtained at barium enema examination shows disappearance of semilunar folds, thumbprinting (arrows), narrowing of the terminal ileum (arrowhead), swelling of the ileocecal valve, and calcifications.

(ii) DIAGNOSTIC IMAGE *showing/revealing* FINDING, for example:

(31) Fig. 1. CT scan of the abdomen showing a large splenic cyst with calcifications in the cystic wall.

The first combinatorial pattern is specific to the language of *Radiology*, whereas the second one only occurs in *European Journal of Internal Medicine.* However, our results showed that the use of one or another pattern is not only related to the type publication, but also to the style of language used and the communicative setting of the publication. Regarding the first combinatorial pattern, all the captions in *Radiology* follow this pattern. Furthermore, they become fixed and institutionalised expressions in the language of *Radiology* since they all follow the same syntactic pattern. The following fragment of KWIC concordance lines obtained from captions in this journal confirms this.

```
1 Figure 4a. Rotational angiograms obtained (a, b) between the posteroanterior and ipsi
2 Posteroanterior chest radiograph obtained at the first admission.  Figure 2. Postero
3   Figure 12. Transverse MR image obtained by using  a diffusion tensor imaging seque
4 vus waveform.  Figure 1b. Images obtained in a patient with a severe right renal arte
5 rmality.      Figure 11b. Images obtained in a young adult  woman with lupus erythema
6 vus waveform.  Figure 1d. Images obtained in a patient with a severe right renal arte
7 .    Figure 2b. Reference images obtained in (a)  control subject and (b) a patient
8  5. Magnification-view mammogram obtained in a patient 4 months after excision of duc
9 y.  Figure 4. Frontal radiograph obtained in a male neonate with right-sided CDH on
10  Figure 1c. Frontal radiographs obtained in neonates with left-sided CDH on the 1st
11 Figure 11c. Transverse CT scans obtained in a 58-year-old woman with Child-Pugh grad
12 Figure 10c. Transverse CT scans obtained in a 77-year-old man with Child-Pugh grade
13 ]        Figure 3a. CT scans obtained in a 72-year-old            man with type 2
14  ion. Figure 5. Blood pool scan obtained with iodinated 131I serum albumin demonstra
```

Figure 2. Fragment of a right-sorted KWIC concordance for obtained (Figure is the context word).

The instructions for submission of manuscripts to *Radiology* are the cause of these fixed and repetitive expressions since they specify that captions should follow the same style used in recent issues of the journal. This means that authors continually copy one another, which results in a very repetitive language. The second pattern only occurs in *European Journal of Internal Medicine* but it does not originate fixed and repetitive sentences as it was the case of *Radiology.* Finally, if we look at the communicative setting, *Radiology* is aimed at just radiologists, whereas *European Journal of Internal Medicine* is aimed not only at specialists but also at a broader readership. This is reflected in the language, which is more formalised and abstract in *Radiology*.

The ECR corpus

Our corpus reveals that authors of ECRs do not normally compose captions for the illustrative figures. Therefore, we are led to focus on the first mechanism for making reference to non-verbal elements.

A.   The text refers to images

The following combinatorial patterns were identified in Spanish ECRs:

(i)   *En el/la* DIAGNOSTIC IMAGE *se observa/se objetiva* FINDING, for example:

(32) En la ecografía se observa una masa sólida de 4 cm de diámetro.

(ii)   *El/La* DIAGNOSTIC IMAGE *muestra/revela* FINDING, for example:

(33) La ecografía abdominal muestra una masa quística de contornos bien definidos.

The first pattern, which includes a passive reflexive structure, is the preferred one in ECRs (77.7%). Figures also show that the Spanish verb *observar* is preferred to *objetivar* in the first pattern, whereas *mostrar* is the verb of choice in the second pattern.
Both patterns match with the ones from PCRs; however, authors of Spanish ECRs do not include the figure number when they make text references to visual elements.
English ECRs show the following pattern.

DIAGNOSTIC   IMAGE   (Fig.   number) shows/showed/reveals/revealed/demonstrates/demonstrated   FINDING, e.g.:

(34) CT Scan (Fig. 1) shows a stone in the distal portion of the right Wharton's Duct.

This combinatorial pattern corresponds to the one from English PCRs, and 'show' is again the verb of choice to introduce the abnormalities found on images. It is used either in the simple present or past.

5.   Discussion and translation implications

The results obtained in this study corroborate our working hypothesis. That is, authors of case reports use different combinatorial patterns to introduce relevant information in the first three moves of the Case Report section. Our study has provided evidence that combinatorial patterns are essential in the development of information in case reports and that these patterns are used according to the communicative aims pursued in this medical genre. Furthermore, both English and Spanish case reports make use of these patterns.

The following table summarises the intralinguistic and interlinguistic equivalences that can be inferred from our phraseological analysis in the PCR and ECR corpora:

MOVES	The PCR corpus	The ECR corpus
Move 1 (presenting the patient's symptoms)	Varón/mujer de ... años de edad con antecedente(s) de/diagnosticado(a) de DISEASE que presenta/acude a/consulta por SYMPTOM ⇔ A …–year-old man/woman/male/female patient was admitted (to) with SYMPTOM	Varón/mujer/paciente de … años con SYMPTOM ⇔ The patient was a (…)-year-old man/woman who presented with SYMPTOM
Move 2 (describing physical examination findings)	A la exploración física el/la paciente presenta(ba)/se aprecia(ba)/se detecta(ba)/destaca(ba) SIGN⇔ Physical examination revealed/disclosed/showed SIGN	En la exploración física se percibe SIGN ⇔ Physical examination revealed/showed SIGN
Move 3 (describing laboratory and other test findings)	The text refers to images: En el/la DIAGNOSTIC IMAGE (Fig. Number) se evidencia/se observa/ se objetiva/se visualiza/se aprecia FINDING ⇔ DIAGNOSTIC IMAGE showed/ revealed/demonstrated/confirmed FINDING (Fig. Number) The caption is included in the image: DIAGNOSTIC IMAGE. FINDING ⇔ DIAGNOSTIC IMAGE (obtained at/in DIAGNOSTIC EXAMINATION) shows/showing/reveals/revealing FINDING	The text refers to images: En el/la DIAGNOSTIC IMAGE se observa/ se objetiva FINDING ⇔ DIAGNOSTIC IMAGE (Fig. Number) shows/reveals/demonstrates FINDING

Table 3: Intralinguistic and interlinguistic equivalences in the PCR and ECR subcorpora.

From a Spanish intralinguistic perspective, brevity is the typical feature in the language of Spanish ECRs. Their combinatorial patterns are much more concise than those of PCRs, sometimes almost telegraphic, with many paratactic structures. As a result, the message gets to the reader faster than in PCRs. English combinatorial patterns in ECRs are not more concise than those of PCRs, but they differ in the degree of formality. Patterns from English journals are more conventional and academic than the ones from e-forums. In addition, the language is more abstract in

Radiology than in *European Journal of Internal Medicine,* due mainly to the former's restricted informative context.

Interlinguistically, table 3 reveals the equivalences established in both languages. Regarding the moves, there are two combinatorial patterns in Spanish printed and electronic case reports to make reference to images in the text, whereas in English case reports there is only one pattern (move 3). On the other hand, authors of Spanish and English electronic case reports do not usually include captions in the images (move 3). Our analysis also confirmed that the captions for illustrations in *Radiology* are fixed imitative expressions which result in a prefabricated language.

As to the verb tenses, Spanish case reports use the simple present to introduce the patient's symptoms, describe test findings and explain the images illustrating abnormalities. In contrast, English always uses the simple past for these three moves. It is our impression that the use of the simple present makes the discourse of Spanish case reports more dynamic, involving the reader in a more active way than the past tense.
In addition, inanimate subjects collocate with active verbs in English case reports ('personification'), which results in clarity and conciseness. The equivalent structures to express this meaning in Spanish are passive reflexive clauses.

Our analysis demonstrated that some discourse strategies of the case report genre are revealed through the combinatorial patterns, so there is "a preferred way of saying things" in the genre as Gledhill puts it (2000: 76). Our analysis also showed that there are prototypical phraseological features specific to the genre and even more specifically to the type of publication. This observation is important because it should make the specialists aware of the type of pattern they should use when writing case reports.

In addition, from a translation perspective, knowing subject-specific phraseological combinations helps to produce a coherent and readable target text since technical translation involves not only understanding the source text and re-formulating it in the target text, but also correctly transposing the terminology and phraseology of the field so as to produce a logically organised target text.

Acknowledgments

Notes

[2] We follow Swales' definition of text genre (1990: 58): "a group of texts that share a set of communicative purposes and are recognized as having legitimacy within a discourse community."

[3] Analysis of the structural organisation of a text genre usually reveals a progression of information divided into *moves*. Nwogu (1997: 122) defines the move as a text segment made up of a bundle of linguistic features which give the segment a uniform orientation and signal the discourse content in it. Each move is taken to embody a number of constituent elements which combine in identifiable ways to constitute the information in the move.

[4] In this paper we use the term *pattern* in the sense of Hunston and Francis (2000:37) since for these authors a pattern can be identified if a combination of words occurs relatively frequently, if it is dependent on a particular word choice, and if there is a clear meaning associated with it.

References

- **Adams Smith**, Diana (1984). "Medical Discourse: Aspects of Author's Comment." *The ESP Journal* 3, 25-36.

- **Baker, Mona** (1987). *A Method for Identifying Statistical Significant Lexis in Medical Journal Articles*. Birmingham: University of Birmingham (unpublished mss).

- **Bayoumi, Ahmed**, M. (2004). "The Storied Case Report." *Canadian Medical Association Journal* 171(6), 669-570.

- **Béjoint, Henri** et Philippe Thoiron (1992). "Macrostructure et microstructure dans un dictionnaire de collocations en langue de spécialité." *Terminologie et Traduction* 2/3, 513-522.

- **Benson, Morton**, Benson, Evelyn & Robert Ilson (1997). *The BBI Dictionary of English Word Combinations.* Amsterdam/Philadelphia: John Benjamins.

- **Bowker, Lynne** and Jennifer Pearson (2002). *Working with Specialized Language. A Practical Guide to Using Corpora*. London/New York: Routledge.

- **Bravo Gozalo**, José María & Purificación Fernández Nistal (1998). "La lingüística del corpus, las nuevas tecnologías de la información y los Estudios de Traducción

en la década de 1990." Fernández Nistal, Purificación & José María Bravo Gozalo (Eds)(1998). *La traducción: Orientaciones Lingüísticas y Culturales.* Valladolid: Universidad de Valladolid, 205-257.

- **Busch-Lauer, Inés-A.** (1998). "Non-verbal elements and data commentary in English Medical Texts." Fortanet, Inmaculada, Posteguillo, Santiago, Palmer, Juan Carlos & Juan Franciso Coll (Eds)(1998). *Genre Studies in English for Academic Purposes*. Castellón de la Plana: Universitat Jaume I, 109-132.

- **Corpas Pastor, Gloria** (1996). *Manual de fraseología española*. Madrid: Gredos.

- **Faber, Pamela** (2001). "Knowledge Representation and Descriptive Terminology Management." Fernández Nistal, Purificación & José María Bravo Gozalo (Eds) (2001), *Pathways of Translation Studies*. Valladolid: Universidad de Valladolid, 147-167.

- **Faber, Pamela** (2002). "Oncoterm: sistema bilingüe de información y recursos oncológicos." Alcina Caudet, Amparo & Silvia Gamero Pérez (Eds) (2002). *La traducción científico-técnica y la terminología en la sociedad de la información*. Castellón: Universitat Jaume I, 177-188.

- **Faber, Pamela** and Maria Isabel Tercedor Sánchez (2001). "Codifying Conceptual Information in Descriptive Terminology Management." *Meta* 46(1), 192-204.

- **Gläser, Rosemarie** (1994/1995). "Relations between Phraseology and Terminology with Special Reference to English." *ALFA* 7/8, 41-60.

- **Gledhill, Christopher** (2000). "The Discourse Function of Collocation in Research Article Introductions." *English for Specific Purposes* 19(2), 115-135.

- **Hunston, Susan** and Gill Francis (2000). *Pattern Grammar. A Corpus-Driven Approach to the Lexical Grammar of English*. Amsterdam/Philadelphia: John Benjamins.

- **Lee, David** (2001). "Genres, Registers, Text Types, Domains and Styles: Clarifying the Concepts and Navigating a Path through the BNC Jungle." *Language Learning and Technology* 5(3), 37-72.

- **Méndez-Cendón, Beatriz** (2005). "Phraseological Patterns in the Language of Medical Case Reports: A Corpus-Based Study." *Corpus Linguistics Conference 2005*. University of Birmingham, United Kingdom.

- **Meyer, Ingrid** & Kristen Mackintosh (1996). "The Corpus from a Terminographer's Viewpoint." *International Journal of Corpus Linguistics* 1(2), 257-285.

- **Motta-Roth, Désirée** (1998). "Discourse Analysis and Academic Book Reviews: A Study of Text and Disciplinary Cultures." Fortanet, Inmaculada, Posteguillo, Santiago, Palmer, Juan Carlos & Juan Francisco Coll (Eds) (1998). *Genre Studies*

in English for Academic Purposes. Castellón de la Plana: Universitat Jaume I, 29-58.

- **Nwogu, Kevin Ngozi** (1997). "The Medical Research Paper: Structure and Functions." *English for Specific Purposes* 16(2), 119-138.

- **Pavel, Sylvia** (1993). "La phraséologie en langue de spécialité. Méthodologie de consignation dans les vocabulaires terminologiques." *Terminologies Nouvelles* 10, 67-82.

- **Roberts, Roda P.** (1994/1995). "Identifying the Phraseology of Languages for Special Purposes." *ALFA* 7/8, 61-73.

- **Roberts, Roda P.** (1998). "Phraseology and Translation." Fernández Nistal, Purificación & José María Bravo Gozalo (Eds.) (1998). *La traducción: Orientaciones Lingüísticas y Culturales*. Valladolid: Universidad de Valladolid, 61-77.

- **Sinclair, John** (1991). *Corpus, Concordance, Collocation*. Oxford: OUP.

- **Swales, John** (1990). *Genre Analysis.* Cambridge: CUP.

- **—** (2004). *Research Genres: Explorations and Applications*. Cambridge: CUP.

- **"Uniform Requirements** for Manuscripts Submitted to Biomedical Journals: Writing and Editing for Biomedical Publication." Online at: <http://www.icmje.org/> (consulted 18.01.2008).

- **Wildsmith, John A.W.** (1995). "How to Write a Case Report." Hall, George M. (Ed) (1995). *How to Write a Paper*. London: BMJ Publishing Group, 64-106.

Biographical note
Beatriz Méndez-Cendón is Senior Lecturer at the University of Valladolid. She teaches English Grammar, Terminology and Translation and Translation of Biomedical Texts. Her principal research interests lie in the areas of medical terminology and phraseology, scientific and technical translation and corpus-based contrastive analysis. She can be reached at <cendon@lia.uva.es>