# LSP Journal

Vol.3, No.1 (2012)

Language for Special Purposes
Professional Communication
Knowledge Management
Cognition

**LSP Journal**

Language for Special Purposes
Professional Communication
Knowledge Management
Cognition

**Vol. 3 No. 1 (2012)**

## Table of contents:

# Editorial

For those who think that knowledge management and cognition are far beyond their daily lives and interests, it may be useful to refer to examples where these issues matter and may even be a question of life and death. The alarming events in Syria in Spring 2012 call for international action. Among several factors which may determine the decision on what the World will do, it matters how the events in Syria are defined and referred to. Is it *unrest, (plain) violence, sectarian violence, a (prolonged) civilian conflict, civil war, all-out/outright civil war, or a non-international armed conflict*? When does a widespread, organized and relentless sectarian conflict qualify as *civil war*? Kofi Annan, the joint United Nations-Arab League envoy to Syria, and a number of observers should provide evidence about the actual events but have had difficulties in getting physical access to specific areas in Syria. Their findings could be used for determining whether it is *sectarian violence* or *civil war* which is a crucial distinction: Defining the events as *civil war* would have repercussions under the international laws of war, e.g. combatants taken prisoners would have to be treated as prisoners of war who cannot be interrogated, nor tortured. However, although the International Committee of the Red Cross could possibly characterize the events as *civil war*, they may refrain from doing so, because of the risk of jeopardizing the humanitarian access negotiated with the Syrian government. The Syrian conflict is indeed a very sad case, where cognition, distinctive definitions, and consistent concept modeling have a crucial role to play.

The LSP Journal has reached 300 subscribers representing a large number of countries all over the World. The journal’s status as a recognized multilingual platform for the publication LSP-related research results is thus becoming more and more manifest, although the number of readers is still relatively modest: 842 unique visitors Jan-May 2012 according to Google Analytics, but this is also a low season given that we publish in June and December.

The call for papers published in our last edition has been fruitful, but we still encourage authors and readers to disseminate the call and at the same time to invite interested colleagues to subscribe to the journal. We would like to underline again that subscription is free and without any constraints.

Henrik Selsøe Sørensen
Editor in chief

http://uk.cbs.dk/staff/henrik_selsoe_soerensen

# Lexical variations in business e-mails written by non-native speakers of English

María Luisa Carrió Pastor [a] and Rut Muñiz Calderón [b]

[a] Universitat Politècnica de València, Spain
lcarrio@idm.upv.es

[b] Universidad Católica de Valencia, Spain
rut.muniz@ucv.es

***Keywords:*** *Lexical variations, e-mails, English as a second language, English as a foreign language*

**Abstract**

English is the lingua franca used in business communication. Therefore the number of non-native speakers of English already outnumbers native speakers provided that worldwide enterprises use English for international communication. The Internet has also increased the use of English as an international language, in this way; it is used by speakers with different linguistic backgrounds. This variety of authors produces differences or variations in language use. In this paper we contrast business e-mails written by Spanish agents who work in an exporting company in India and China. Our main aim is to analyze the possible variations due to the mother tongue and the socio-cultural context and to classify lexical variation in business English used as a global working language by non-native speakers. We intended to determine the causes of variation and their influence on discourse. We analyzed and contrasted sixty e-mails written by two groups of non-native English speakers. Group A was composed of native speakers from Pakistan and Group B was composed of native speakers from China. The corpus analysis was carried out manually. We classified the occurrences in categories depending on the cause of the variation. After the analysis, we observed that the lexical variations found were caused by sociolinguistic and cultural influences.

## 1 Introduction

There is no doubt nowadays that English is the lingua franca for the global business world. Europe has seen the growth of English and, as a consequence, it has become the dominant language in the European Union (Bargiela-Chiappini and Nickerson, 2002). In Asia, the English language is used as a lingua franca for the international relations, although the Chinese language is gaining importance being China one of the greatest economic potencies in the world.

In this study, our main aim is to classify the lexical variations that may be found when contrasting the language used by non-native speakers of English. In this sense, we have chosen a genre in which communication is crucial: a business environment. We have also considered interesting to analyse the expressions used in electronic communication. The immediacy of messages in business contexts could provide variations in the use of the English language as writers may not revise the message being sent. In e-mail communication linguistic and cultural backgrounds are more evident as the main intention of the writers is to take decisions or send orders as quickly as possible (Crystal, 2005; Kessler, 2010). In this way, after the contrast of the e-mails written by business workers from Pakistan and China, variations were detected and classified in order to show that standard English is not a reality in a multicultural and ever changing world.

We have also considered that the widespread of the English language has been caused mainly by two facts: the colonization of the British Empire from the seventeenth to the twenty-first century in Africa and Asia, and the strong American influence in the finance and commerce contexts during the twentieth century. Therefore, all over the world, people use more and more varieties of English in ever-increasing numbers (Kirkpatrick, 2007). These different English varieties that have flourished along the years have their own features. They share history and their affinity with either British or American English, but there are also some aspects that are unique to each variety, particularly in terms of accents, in their idiomatic use of vocabulary, their application of grammar and their discourses strategies (Jenkins, 2006).

Another important factor is that technology changes the way business workers are in contact and this new communication environment implies changes in language use as well as changes in the business trade itself (Crystal, 2005; Mateo, 2007: 196). More precisely, the Internet has revolutionised the way enterprises communicate. The communication is immediate and fast, the style is more direct, the discourse is presented in a simple and dense way and an instant answer is frequently expected. Furthermore, there are not linguistic barriers; i.e. people with different linguistic backgrounds communicate in English that is considered as a lingua franca.

International business discourse, as any type of discourse, is culturally-situated and, therefore, context-dependent. All these three aspects, i.e. discourse, culture and context, play a key role in the communication process (Bargiela-Chiappini and Nickerson, 2002: 31-34). The use of the English language by international writers in business' exchange of ideas is a phenomenon that causes changes in writing even in those texts which are most expected to follow standard guidelines (Crossley and McNamara, 2009; Carrió and Muñiz, 2010; Ädel and Erman, 2012).

Despite the fact that English is a lingua franca in the business and trade contexts we have also to take into account that writers express in different ways even in a standardised context as business. Writers from different backgrounds write in different ways and variations can be found due to their cultural and linguistic antecedents. Variation is considered in this study as the differences in the synchronic use of a language due to several causes: interlanguage (Ellis, 1999), rhetorical aspects (Samraj and Monk, 2008; Durrant, 2009), genre variation (Charles, 2007; Ozturk, 2007), and the use of language by native and non-native speakers (de Haan & van Esch, 2005; Moreno & Suárez, 2008; Hinkel, 2009; Schleef, 2009). In this study we analyse the last aspect, variation caused by the different linguistic background of the speakers.

We should also take into account that variation in intercultural communication has increased due mainly to second language and culture learning (Alptekin, 2002) and global communication technologies (Thatcher, 2005). Some researchers have also analysed the different ways of communication through e-mails in a business context (Nickerson, 2005; Waldvogel, 2005, 2007; Kankaaranta, 2006; Jensen, 2009). Nevertheless, we should highlight the fact that there is still a lack of research in the specific area of English intralanguage variation related to international business communication (Rivers, 2008).

In this study, our analysis is focused in finding the lexical variations in business e-mails written by non-native writers. We focus our research in business English, as it is a genre with lexical patterns internationally established in order to communicate fluently. Business English lexis is a structured and relatively homogenous whole. Little appears in the lexis of business English by chance and it is framed by an infrastructure that grows out from the level of the word through collocation, clustering, colligation, semantic prosody, genre and on to a supra-generic ordering, until it meets general English. It is relevant for businessmen to communicate fluently with companies all around the world and this is the reason why the vocabulary tends to follow standard patterns. Nonetheless, in this study we would like to confirm if cultural or mother tongue influences manipulate the way international writers communicate.

In this paper, we analysed if some features of communication change due to the mother tongue and the socio-cultural context of the writers in virtual business. Our main objective is to classify the lexical aspects that change. A further objective is to detect the lexical variations in those that are caused by the mother tongue influence or by the sociolinguistic and the cultural context of communication.

## 2 Methodology

### 2.1 The corpus of the research

In order to extract our results, this study used a corpus of authentic sixty e-mails written by non-native English speakers with different cultural and linguistic backgrounds. One group speaks English as a foreign language (texts from China) and the other group speaks English as a second language because it is a former British colony (texts from Pakistan). These e-mails were all written during a period of two years, from 2009 to 2011, by the Chinese and Pakistani employees of a Spanish exporting company. Due to privacy concerns, this corpus of real life e-mails was not constructed for public use; it was only used for this study.

This enterprise is specialized in the manufacturing of high technology laser machines, which are used for the finishing of denim. Moreover, one of the main activities of this enterprise is selling and exporting this equipment worldwide. Nowadays, many American and European manufacturing companies have outsourced their productivity to the Southeast of Asia. Therefore, this company's major customers are in this part of the world, where Pakistan and Chinese buyers are its most important clients. The Sales Department is managed by the Sales Manager, who supervises the negotiations and agents located in different countries dealing with customers worldwide. Presently, the area of most commercial activity for this enterprise is the Asia-Pacific region, where the cost of labour is lower than in other parts of the world.

English is the language used for business communication and the most usual way of exchange of ideas inside and outside the company is through e-mails, which are sent through a laptop or smartphone, which allows communication to be almost instant. This way of communication is very useful for the business professionals under study who are out of the office most of the time.

The majority of the e-mails compiled in the corpus are usually sent or received within the Sales Department. All the participants are non-native English speakers although their lingua franca is English. In addition, the use of English is compulsory even among workers with the same mother tongue, so that everyone can communicate fluently.

We noticed, after a general reading of the corpus, that the sociolinguistic background of the writers of the e-mails was very different. They are all native speakers of Mandarin and Urdu. However, their linguistic competence in English varies due to the way they acquired English. For instance, the Chinese writers learnt English as a foreign language, as an optional subject in their educational system; meanwhile, the Pakistani writers learnt English as a second language, and it has a central position in their national life (Kachru, 2005). The role of the English language in Pakistan has been an essential one ever since the foundation of the country in 1947. It is used in the contexts of education, economy, business, political life, mass media, science and technology. English is the language of communication at the level of trade and commerce as well as in the global industries. Therefore, it plays an important function not only in the business context itself but also in the country's development (Tahir, 2004).

As for the situation of the English language in China we can add that after the official access of China in the World Trade organization in 2001, the Chinese business and trading communities adapt to international markets and the popularity of English seems to have reached a new position with government policy-makers, educationalists and the Chinese public (Bolton, 2003).

### 2.2 The analysis of the corpus

The corpus of sixty e-mails was selected from a larger corpus of one hundred and twenty e-mails collected from the exporting company. The e-mails were selected considering the mother tongue of the sender of the message. The managers of the company receive mails from workers from different Eastern countries, such as India, China, Pakistan, etc. The e-mails included various subjects such as invitations to fairs, arrangements for trips, sales, etc.

The analysis was carried out manually because our main purpose was to detect the lexical variations in context so we preferred to select the occurrences considering their syntactic and semantic nature. The analysis was carried out in several steps. Firstly, all the e-mails were carefully read and were classified into two different groups according to the writers' origins. Secondly, the lexical variations found were noted down and, later they were classified in the different causes of variation. Finally, the total number was calculated as well as the percentages of use.

After this, we contrasted the results, which were classified in tables, and, finally, we explained the conclusions of this research. This study of lexical variations in business e-mails is one aspect of a larger research study, which is focused on the analysis of the computer mediated communication in a workplace context among non-native English speakers with different cultural backgrounds.

## 3. Results and discussion

We mainly found two classifications for the lexical variation in the corpus: noun phrase clusters and discourse style. As it is shown in Table 1, the total occurrences of the variation in the use of noun phrases detected in the thirty e-mails written by Pakistani workers were 15. The compound words made by an adjective and a noun (ADJ+N) had a higher frequency of variation (40%), followed by the compound nouns made by two nouns (33%). In the thirty e-mails written by the Chinese businessmen we obtained more variations in the different categories. We observed that the use of invented names, i.e. the writers invented names to adapt to clients.

| Noun phrase clusters | Pakistan occurrences (percentages) | China occurrences (percentages) |
|---|---|---|
| N+N | 5 (33%) | 4 (11%) |
| N+N+N | 2 (13%) | 4 (11%) |
| N+ADJ+N+N | - | 2 (6%) |
| N+N+N+N | - | 1 (3%) |
| N+PREP+N | 1 (7%) | - |
| ADJ+N | 6 (40%) | - |
| First name invention | - | 24 (69%) |
| ADJ+N+N | 1 (7%) | - |
| TOTAL | 15 (100%) | 35 (100%) |

**Table 1.** Noun phrase clusters in e-mails from Pakistan and China.

The non-native writers varied the writing of most of the noun clusters as a consequence of their mother tongue influence. They imitated the way they express in their mother tongue and changed their way of communication. There is one exception to this cause, the invention of first names. This was caused by a sociolinguistic influence. They imitated the European names in order to adapt their communication to the receivers of the message. We could see some of the variations found in the Pakistani e-mails and the implication of variation, they used different ways to communicate as we can see in Example 1:

Example 1 :

N+N: chiller liquid.
ADJ+N: foreign principals.
N+N+N: family laser company; opening demo Center.

In the e-mails written by Chinese businessmen we can see examples of different ways to use noun phrase clusters and first name invention as we can see in Example 2:

Example 2 :

| |
|---|
| N+N+N+N: Chinese New Year Holiday, Laser company China office. |
| N+N: China factory, Chinese boarder. |
| N+N+N: China custom clearance, Hong Kong and China clients, Asia countries. |
| First name invention: Cason, Florence, Williams, Winney. |

The highest number of occurrences found in the Chinese group refers to the invention of first names. Chinese writers choose an occidental name to sign their e-mails in order to facilitate the European or American readers the comprehension of their names, because Chinese names are usually very difficult to pronounce. This is a very common practice that has cultural connotations because Chinese people are very respectful and considerate so this is a way of showing it.

In Table 2 the occurrences due to several discourse variations in which lexical elements are involved. We grouped these variations under the label discourse as they affect several aspects:

| **Discourse variations** | **Pakistan occurrences (percentages)** | **China occurrences (percentages)** |
|---|---|---|
| Stylistic ornamentation and politeness | 5 (28%) | 5 (10%) |
| Lexical calques | 4 (22%) | 2 (4%) |
| Abbreviations | 5 (28%) | 30 (62%) |
| Exaggerations | 2 (11%) | 0 (0%) |
| Reduplication | 2 (11%) | 0 (0%) |
| Direct and imperative style | 0 (0%) | 10 (20%) |
| L1 adaptation | 0 (0%) | 2 (4%) |
| **TOTAL** | 18 (100%) | 49 (100%) |

**Table 2.** Discourse variation in e-mails from Pakistan and China.

Pakistani writers do not vary the way they express such much as the Chinese writers. Pakistani businessmen varied the way they abbreviate and the stylistic ornamentation and calques (28%). Chinese businessmen vary in the way they use abbreviations (62%) and in the use of direct and imperative language (20%). The occurrences found due to the stylistic ornamentation, the lexical calques, the exaggeration, the direct and imperative style and the L1 adaptation were caused by the influence of the mother tongue of the writers. Other occurrences found in this category, as abbreviations and reduplications were caused by a sociolinguistic cause, the business environment of the writers changed the way they communicate. We can observe several examples of the occurrences found in the e-mails from Pakistan in Example 3:

Example 3 :

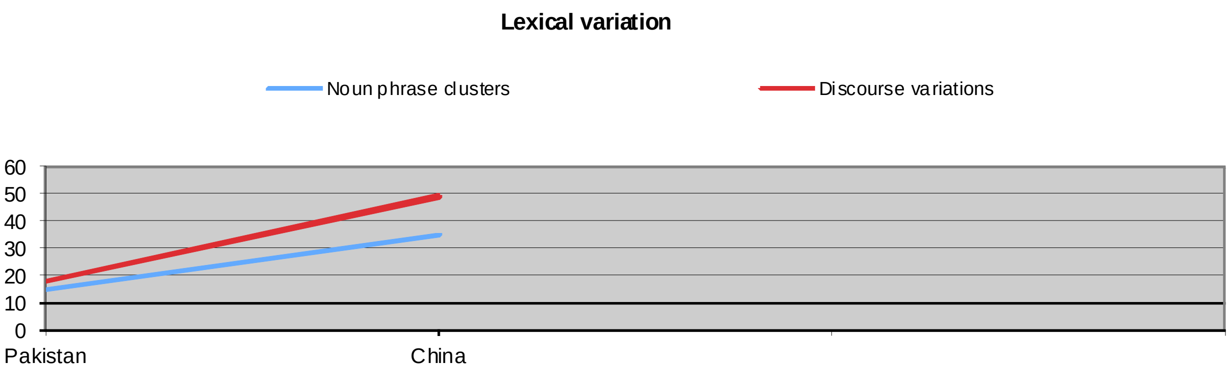
| |
|---|
| Lexical calques: In the largest interest of Laser machine company, I am pleased to extend the hand… We need to attend the problem. |
| Stylistic ornamentation and politeness: In the largest interest of Laser machine company, I am pleased to extend the hand…, kindly select the space from the floor plans, I would like to bring to your notice my disappointment. |
| Abbreviations: Please allow me to tell u his type… |
| Reduplication: ...as well after many many meetings…. I have repeatedly called several times. |
| Exaggerations: I am the only one in 200 Million population of Pakistan who knows laser machine, He fell in love with the samples of CDS and this going to open a new market to us, He is a poker player and a master of bluffing. |

The examples we found in the Chinese e-mail are ones shown in Example 4:

Example 4 :

| |
|---|
| Lexical calques: No need to make it worry, the behind message… |
| L1 adaptation: Don't leave this argument between you and me, This kind of argument do not bring a harmony business relation, My concern is more deep in behind. |
| Abbreviations: Pls find out and…, your proposal is not a solution needed to wait for after CNY holiday, Pls advise yr fty exact address, tks, your G2 rsult samples. |
| Politeness expressions: Please Kindly. |
| Direct and imperative style: We want all the parts sent us with the next Big M. shipment together, We have no room for Big m., it's not make sense to have room for production, Ok, we have tried everything now. I think its time to get some one here R., you are already in china, come to ningbo and sort this out, I want to fix the work schedule as well. |

In Graph 1 we contrasted the results obtained in the two different classifications of lexical variations:




**Graph 1.** Contrast of lexical variation in the e-mails written by non-native English speakers.

As it can be observed, the variations produced by non-native English speakers with a high level of this second language are quite evident. They change the way they communicate. We can observe that the variations produced by the second language writers (Pakistani) are not frequent; nevertheless, the variations produced by the foreign language writers (Chinese) are higher. We also observed that most of the variations occurred due to the influence of the mother tongue, although some other variations were caused by sociolinguistic causes, which were considered in this study due to the business environment of the writers. They are forced to write in a certain way but they do not use standard English, they vary the way they communicate. This fact shows that the English language changes and varies depending on the genre but also on the communicator.

**4. Conclusions**

When we contrasted the two groups, we found that the group of e-mails from China showed a higher percentage in complex noun phrases than the group of Pakistan. The reason for this is the clear influence of the mother tongue in China, compounding is very commonly used among Chinese when speaking their first language as well. The speakers tend to create new words imitating their mother tongue, as they need to name what they do not know; even when English offers a lexical solution for this, the Chinese speakers of English choose to form a new word. We also observed in the answers to these e-mails that their interlocutors understood them; the variation in the use of the English language is not an obstacle for communication.

When contrasting the rest of the occurrences found in the discourse variation between the two groups of writers, we observed that the lexical feature that showed more cultural influence in the Chinese group of writers was the invention of first names and the writing style. However, the style used by the Pakistani writers is more neutral even though we have detected some ornamented style expressions with some cultural influence. Business English writing tends to be concise and more direct so this group adapts a more standard business style than the Chinese.

Furthermore, the Chinese group used a direct and imperative writing style. This way of writing is very influenced by the Chinese language, especially Mandarin. The overall style tended to be brief and direct with the frequent use of the modal verb *must* that is considered imperative. This was a socio-cultural influence because Chinese languages are very direct, the style is synthetic, modal verbs are hardly ever used; therefore the use of this modal might sound too commanding for the European and American counterparts. Finally, it was observed that the fact that Pakistani writers use English as a second language, i.e. English is considered relevant in their society, improved the way they communicate, varying less the way they express. On the contrary, Chinese writers, who learnt English as a foreign language, vary the use of English; they suffer interference in their linguistic production. In this sense, language contact is a valuable asset to be considered.

This study could not evaluate all the aspects related to lexical variations, we only focused on some aspects, further research should be carried out in order to classify and understand the causes of the variations that can be observed when writers with different cultural and linguistic background use a second or foreign language. Furthermore, some outcomes form this research might also be topical subjects for further studies to analyse social media and the implications of English as the lingua franca in a globalised world.

## 5 References


Ädel, A. and Erman, B. (2012): Recurrent word combination in academic writing by native and non-native speakers of English: A lexical bundles approach. *English for Specific Purposes*, 31, 81-92.

Alptekin, C. (2002): Towards intercultural communicative competence in ELT. *ELT Journal*, 56(1), 57-64.

Bargiela-Chiappini, F. and Nickerson, C. (2002): Business Discourse. *International Review of Linguistics in Language Teaching*, 40- 4.

Bolton, K. (2003): *Chinese Englishes. A Sociolinguistic History.* Cambridge: Cambridge University Press.

Carrió, M. L. and Muñiz, R. (2010): Variations in business English letters written by non-native writers. *LFE. Revista de Lenguas para Fines Específicos*, 16, 39-56.

Charles, M. (2007): Argument or evidence? Disciplinary variation in the use of the noun that pattern in stance construction. *English for Specific Purposes*, 26, 203-218.

Crossley, S. A. and McNamara, D. S. (2009): Computational Assessment of lexical differences in L1 and L2 writing. *Journal of Second Language Writing*, 18, 119-135.

Crystal, D. (2005): *La revolución del lenguaje*. Madrid: Alianza.

Durrant, P. (2009): Investigating the viability of a collocation list for students of English for academic purposes. *English for Specific Purposes*, 28, 157-169.

Ellis, R. (1999): Item versus System Learning: Explaining Free Variation. *Applied Linguistics* 20(4), 460-480.

Haan, P. de & K. van Esch. (2005): The development of writing in English and Spanish as foreign languages. *Assessing Writing*, 10, 100-116.

Hinkel, E. (2009): The effects of essay topics on modal verb uses in L1 and L2 academic writing. *Journal of Pragmatics*, 41, 667-683.

Jenkins, J. (2006): Current perspectives on teaching World Englishes and English as a Lingua Franca. *Tesol Quaterly*, 40(1), 157-181.

Jensen, A. (2009): Discourse strategies in professional e-mail negotiation: a case study. *English for Specific Purposes*, 28, 4-18.

Kachru, B. (2005): *Asian Englishes today. Asian Englishes beyond the canon.* Hong Kong: Hong Kong University Press.

Kankaanranta, A. (2006): "Hey Seppo, Could You Please Comment on This!" Internal Email Communication in Lingua Franca English in a Multinational Company. *Business Communication Quarterly*, 69, 216-224.

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

Kirkpatrick, A. (2007): *World Englishes. Implications for International Communication and English Language Teaching.* Cambridge: Cambridge University  Press.

Mateo, J. (2007): El lenguaje de las Ciencias Económicas. In E. Alcaraz Varó, J. Mateo Martínez and F. Yus Ramos (Eds.) *Las lenguas profesionales y académicas.* Barcelona: Ariel.

Moreno, A. I. & L. Suárez. (2008): A study of critical attitude across English and Spanish academic book reviews. *Journal of English for Academic Purposes*, 7, 15-26.

Nickerson, C. (2005): "English as a lingua franca in International Business contexts." *English for Specific Purposes*, 24, 367-380.

Ozturk I. (2007): The textual organisation of research article introductions in applied linguistics: variability within a single discipline. *English for Specific Purposes*, 26: 25-38.

Rivers, D. J. (2008): English as an International Business Language (EIBL): The Need for and Increase in Theoretical and Practical Research Focusing on Written Business Communication across Cultural Boundaries in Relation to Multinational Corporate Language Selection. *Asian ESP Journal*, 4, 6-22.

Samraj, B. & L. Monk. (2008): The statement of purpose in graduate program applications: Genre structure and disciplinary variation. *English for Specific Purposes*, 27, 193-211.

Schleef, E. (2009): A cross-cultural investigation of German and American academic style. *Journal of Pragmatics*, 41, 1104-1124.

Tahir, A. (2004): Business English: role and implications in Pakistan. In Sabiha Mansoor, Shaheen Meraj y Aliya Tahir (Eds.), *Language policy, planning and practice: A South Asian perspective* (pp. 205-211). Karachi: Oxford University Press.

Thatcher, B. (2005): Situating L2 writing in global communication technologies. *Computers and Composition*, 22, 279-295.

Waldvogel, J. (2005): *The Role, Status and Style in the Workplace Email: A Study Zealand Workplaces*. Doctoral Dissertation. School of Linguistics Applied Language Studies. New Zealand: Victoria University of Wellington.

Waldvogel, J. (2007): Greetings and Closings in Workplace Email. *Journal of Computer-Mediated Communication*, 12, 456-477.

***

# Reusing lexical resources in the construction of a social science multilingual thesaurus

Chryssa Kappi [a] and Lorna Balkan [b]

[a] Institute for Social Policy
National Centre for Social Research (EKKE)
Athens, Greece
ckappi@ekke.gr

[b] UK Data Archive
University of Essex
Colchester, UK
balka@essex.ac.uk

***Keywords:*** *Social sciences, lexical resources, multilingual thesauri*

**Abstract**

We describe how existing lexical resources were used in the development of the social science multilingual thesaurus European Language Social Science Thesaurus (ELSST) which was designed to be used for indexing and finding research data and metadata. The lexical resources of particular interest are *classifications, terminologies, thesauri,* and *controlled vocabularies*, collectively known as *Knowledge Organisation Systems* (KOS). We first describe the properties of these different kinds of resources, which were designed for different purposes and audiences, and then discuss specific examples of classifications, terminologies and thesauri that were used in the development of ELSST. We propose possible solutions to problems we encountered that could enhance the reusability of these resources.

## 1 Introduction

Lexical resources are increasingly important tools for managing and retrieving data and information on the web. There is however a plethora of resources available, most of which have cost a lot of time and effort to produce. Much research effort, therefore, in recent years has gone into exploring ways in which to make these tools reusable and interoperable (van Hage et al., 2008; Lauser et al., 2008).

Here we describe how a variety of lexical resources that were designed for different purposes and audiences were used in the development of the social science multilingual thesaurus European Language Social Science Thesaurus (ELSST), and discuss some problems we encountered.

The lexical resources we looked at are knowledge organisation systems (KOS)[1]. Zeng and Hodge (2011)define knowledge organisation systems as "*all types of schemes for organizing information and promoting knowledge management…Different families of KOS, including thesauri, classification schemes, subject heading systems and taxonomies, are widely recognised and applied in both modern and traditional information systems*." In this paper we are concerned with KOS that are most relevant to archiving research data and metadata in the domain of social sciences, namely *classifications, terminologies* and *thesauri*. Most of these resources are controlled vocabularies, i.e. restricted sets of terms, as opposed to natural language, where there is no restriction on vocabulary use. Controlled vocabularies are intended to facilitate communication by resolving two problems associated with natural language, namely homography where two or more words have the same spelling, but mean different things, and synonymy, where two or more words have identical or similar meanings.

Lexical resources have always played an important role in data archiving and multidisciplinary and cross-national research in the social sciences. Prior to the construction of ELSST, European social science data archives had to rely on locally developed indexing tools including monolingual thesauri. The coverage of ELSST was designed to reflect the content of existing collections in European social science data archives, which means that aside from being multilingual, it was also desirable that it be compatible and interoperable with:

a) international standard classifications
b) social science documentation standards and relevant documentation software (e.g.: Data Documentation Initiative (DDI) (DDI Alliance, 2009))
c) other related controlled vocabularies
d) terminology in other related scientific fields (e.g. statistical terminology)

## 2 Short history of the construction of ELSST

ELSST has been developed over the last decade by the members of the Council of European Social Science Data Archives (CESSDA) to support access to its data catalogue. Data collections span sociological surveys, election studies, longitudinal studies, opinion polls, and census data, including international and European data such as the European Social Survey, the Eurobarometers, and the International Social Survey Programme. The development of ELSST was funded by a series of international projects[2] which aimed to create new or adapt existing controlled vocabularies for the representation of data and metadata in social sciences archives and other data organisations (Figure 1).

ELSST was originally developed from the English monolingual Humanities and Social Science Electronic Thesaurus (HASSET) created by the UK Data Archive at the University of Essex[3].The process of adapting HASSET for ELSST was dictated by its multilingual character and its functionality in relation to comparative research, thus it involved stripping out lower level culture-specific terms and reworking term hierarchies, to make them more

[1] This term originated from the Networked Knowledge Organisation Systems Working Group at its initial meeting at the ACM Digital Libraries '98 Conference in Pittsburgh, Pennsylvania according to Hodge (Hodge, 2000).
[2] For example the LIMBER, MADIERA and CESSDA PPP projects (Miller & Matthews, 2001; Alvheim, 2006; CESSDA PPP, 2008-2010).
[3] Currently ELSST is being used for searching CESSDA collections; It is available for the general public to view at the following web page: http://elsst.esds.ac.uk/login.aspx. Anyone wishing more information on ELSST should contact Lucy Bell: lajbell@essex.ac.uk

relevant across cultures. In the latter phase of its construction, as emerging CESSDA and other European projects were increasingly aiming at standardisation of research and archiving methods in comparative research, the need to make ELSST interoperable with co-existing relevant lexical resources led to its being reviewed and restructured(CESSDA PPP, 2008-2010). All work, including the search for other language equivalents, was undertaken by a multilingual, multidisciplinary team, consisting of language experts and subject specialists.

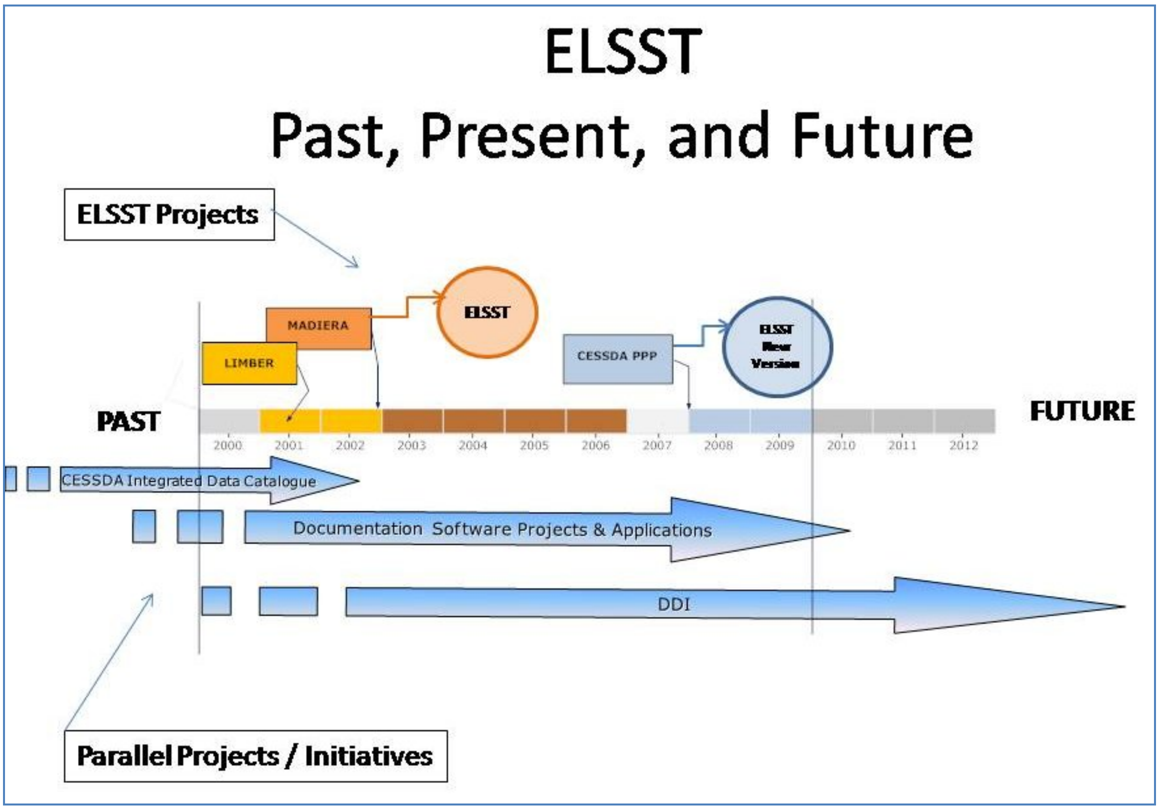



**Figure 1.** ELSST timeline (this figure is an adaptation from Miller (2004))

ELSST is currently available in nine languages – Danish, English, Finnish, French, German, Greek, Norwegian, Spanish and Swedish, with over 3,000 terms for the majority of its languages. Users can browse the thesaurus, perform free-text or keyword searches in their preferred language and retrieve all relevant resources not only in that language but in any of the nine supported languages. In this way, ELSST promotes cross-national comparative research in the social sciences.

## 3 Overview of lexical resources

In this paper, we focus on knowledge organisation systems (KOS), which are key resources for information retrieval. There is no one definition of KOS, although it is generally agreed that they include at least the following subtypes, as defined by Hodge "*term lists, which emphasize lists of terms often with definitions; classifications and categories, which emphasize the creation of subject sets; and relationship lists, which emphasize the connections between terms and concepts.*" We show this schematically in Figure 2, where we have added terminologies to Hodge's category 'term lists'.

We have also added controlled vocabularies to Hodge's classification. Thesauri are a type of controlled vocabulary. Controlled vocabularies resolve ambiguity by adding qualifiers to homographs. For example in ELSST, the medical sense of ‘labour’ is distinguished from the work-related sense by the addition of a qualifier in brackets: LABOUR (PREGNANCY) and LABOUR (WORK) respectively. Synonymy in controlled vocabularies is handled by identifying one term as the preferred or controlled term and listing terms with identical or similar meanings as synonyms.

Figure 2 is intended to show that controlled vocabularies can span each category of KOS.

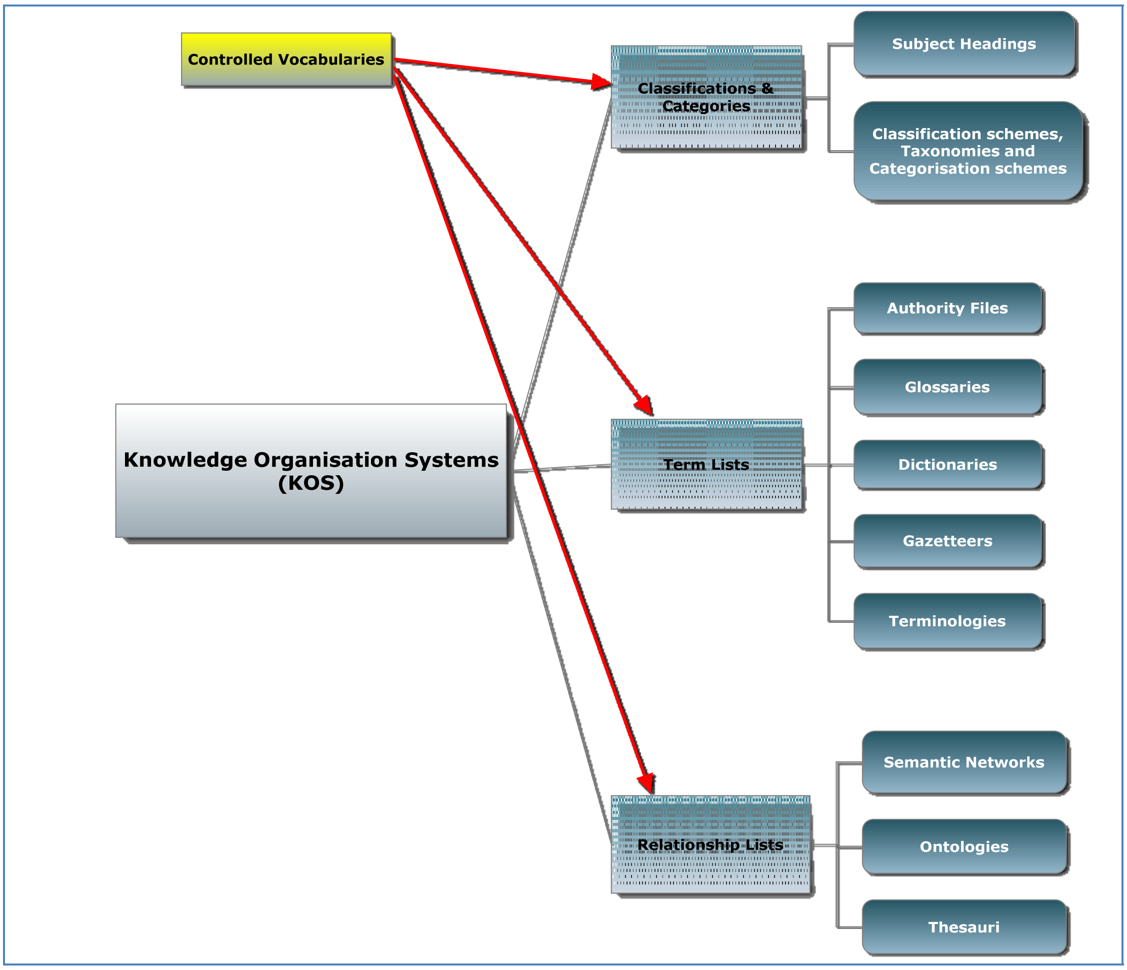



**Figure 2.** A classification of Knowledge Organisation Systems (KOS)

## 3.1 Lexical resources: definitions

### *3.1.1 Classifications*

Classification is the act of any grouping of items under specific organising principles or criteria (classifiers). It is defined by Hayek (Hayek, 1976) as a process “*in which on each occasion on which a certain recurring event happens it produces the same specific effect, and where the effects produced by any one kind of such events may be either the same or different from those which any other kind of event produces in a similar manner […]*”. Classifications are the product of the classification process and can be described as a set of terms representing concepts or classes, listed hierarchically or in some other systematic fashion in order to display relationships among them.

Classifications are the oldest type of KOS, in the sense that they predate the internet; they can be traced back to the first attempts at organising scholarly documents, serving the communication needs of the knowledge community –"*the Republic of Science*", as Guédon calls it (Guédon, 2008). This led to the production of the early classification systems, such as the *Dewey Decimal Classification* (DDC) (DDC, 1989), first published in 1876, and the *Library of Congress Classification* (LCC) (The Library of Congress, 2009).Early classification systems served to promote communication between two parties: indexer and information retriever. For early 20th century libraries this was a rational model, because the librarian was meant to serve the academics and researchers across all scientific disciplines without being an active agent of the research process. Nowadays the indexer tasked with classifying objects faces greater challenges, including the vast amount of information available, its exponential rate of growth, and the trend towards ever greater specialisation and integration of knowledge made possible by technology (Kappi, 2009). Problems are compounded by the fact that the information retriever is often not a trained librarian, but more often a primary researcher, unlike the early 20th century model.

### *3.1.2 Terminologies*

ISO 1087 (ISO, 1990) defines terminology as a *"set of designations belonging to one special language"*, where designations can be terms, appellations or symbols.A terminology usually includes definitions of the designated objects (i.e.: concepts) and possibly other information such as sources and examples. While in the past many terminologies were word-based, the more common method nowadays is to start from the concept.

### *3.1.3 Thesauri*

A thesaurus consists of a structured set of terms. Terms can be either descriptors, i.e. controlled terms that are used for indexing purposes, or non-descriptors (or synonyms), which cannot be used for indexing purposes but which act as pointers to the descriptors. Descriptors denote one and only one concept. Terms are related via a fixed set of relationships which typically include hierarchy (represented by the notation BT (broader term) and NT (narrower term)), synonymy (UF (Used for)) and association or relatedness (RT (related term)). Scope notes (SN) are also provided if necessary to clarify meanings or usage of a term. International standards exist for thesaurus construction (see, for example (ISO, 1985; ISO, 1986)).

| **Characteristics** | **Lexical Resources** | | | |
|---|---|---|---|---|
| | **Classifications** | **Thesauri** | **Terminologies** | **Controlled Vocabularies** |
| **Definitions** | Optional | Optional | Optional | Optional |
| **Relations Schema** | Optional | Necessary | Optional | Optional |
| **Domain Specificity** | Optional | Optional | Necessary | Optional |
| **Function** | Categorizing objects with common characteristics | Organising terms and defining their relationships for indexing and information retrieval | Promoting communication in specific knowledge domains | Controlling use of terms in disparate information sources |

**Table 1**. Lexical resources: a comparison

Table 1 above presents a comparison of lexical resources based on a set of key characteristics. There is, however, no hard and fast distinction between the different types of lexical resources. Thus, for example, classifications may simply provide a structured view of a set of objects without any definition of concepts and relations; or they may take the form of thesauri where concepts and relations are defined. Likewise, terminologies with hierarchical structures can be considered highly specialised classifications; and a controlled vocabulary *"[…] can range from a short list of clearly defined, mutually exclusive, and exhaustive terms, which are the only choices for usage in a specific context […] through a classification to something as complex as a thesaurus with thousands of terms and term relationships."*(Granda, Kramer & Linnerud, 2008).

**3.2 Lexical resources: functions**

Table 1 shows the main functions of each of the above lexical resources. However, the different types of resources may be regarded as complementary to each other, serving different purposes. For example, in medical practice, clinical terminologies are used for specific and detailed annotations of cases, while classifications may be used for organisational and administrative record-keeping for the care provided (Campbell & Giannangelo, 2007). Difference in function of lexical resources is reflected in differences in their content and structure. For example, terminologies tend to be much more fine-grained than classification systems, since the whole point of the latter is to provide broader ways of classifying objects.

All the above lexical resources help to promote consistency in language use within (and possibly beyond) the organisation in which they were created, and thus improve communication. Increasingly, as they become more amenable to machine interpretation they are also finding new uses in computer applications such as text mining. Controlled vocabularies in particular are set to play a critical role in the semantic web.

Multilingual lexical resources have an additional role as support to translators, and as tools to allow access to data across geographical or cultural divides. As the Semantic Web grows, so too does the demand for multilingual accessing and querying of knowledge repositories and linked data that are becoming available.

There is also an increasing trend towards standardisation of research data and metadata (including controlled vocabularies) worldwide. Examples include the DDI (DDI Alliance, 2009) an international alliance which works on a number of relevant projects, and the existence of international standard classification systems, such as International Standard Classifications (ISC) for specific subject domains, e.g.: International Standard Classification of Occupations(ISCO)(ILO, 2011), International Classification of Functioning, Disability and Health(ICF) (WHO,2010b),International Classification of Diseases(ICD) (WHO,2010a), International Standard Classification of Education(ISCED) (UNESCO, 2006), etc.

**4 Concept analysis as foundation for building lexical resources**

**4.1 Overview**

Common to most of the lexical resources discussed in Section 3 above, is the need for conceptual analysis to achieve high quality. By conceptual analysis we mean the breaking down of the concepts represented by the linguistic labels or terms into their component parts to better understand and/or define them.

Concept analysis is particularly important for establishing multilingual equivalence between terms, since the meaning of concepts may be culture-dependent.

While concept analysis is acknowledged as necessary for high quality lexical resources, there is no agreed methodology for how to do it. All approaches to concept analysis, however, start with three steps, regardless of the knowledge domain: (1) define the goal(s), (2) define the scope, i.e. the domain boundaries and (3) list the concepts to be analysed (see Nuopponen, 2010a; Nuopponen, 2010b; Nuopponen, 2011).The essential aim of concept analysis, irrespective of motivation and method, is the *transfer of meaning.* This aim is of particular importance in the social sciences domain since meaning is an essential element in any communication act (see for example Chouliaraki & Fairclough, 1999; Harbsmeier, 2007). Guarino (1995) argues that to accomplish the transfer of meaning, it is not enough to identify concepts and their designation -the relationships between concepts must also be defined. Context analysis also involves examining the context in which concepts are used (e.g. sources documents) in order to identify the key characteristics which help to define and differentiate them. The process requires the skills of both linguists and domain specialists (see for example ISO, 2009;Priss, 1996; Priss, 2006).

### 4.2 Concept analysis for lexical resources in the social sciences

Social science terms have a number of characteristics that make them ideal candidates for concept analysis. Firstly, many social science terms are intrinsically vague or ambiguous, even at a monolingual level. This is because there is a large overlap with general language.

Secondly, there is a tendency, particularly in applied social sciences, to treat terms as though they express invariant concepts which can be used and re-used for research purposes irrespectively of temporal, spatial, social or other context. However, the meaning of social science terms, like any other words of a language, can evolve over time, as a result of societal changes. For example, what ‘old age’ denotes today is not the same as it denoted 20 years ago, and what it possibly might denote in another 20 years, as life expectancy rises.

This creates problems for the design, analysis, preservation and dissemination of research data. In field research, for example, where concept definitions do not form part of the documentation, variables and units of measurement are assigned linguistic labels, i.e. terms, which ‘imply’ the meaning of underlying concepts, without defining them. Examples include ‘childhood’, ‘household’ and ‘special needs’. Thus, ideally, any type of social study should be accompanied by a formal concept analysis process for its study objects. In practice, concepts are only defined where required by the study, as in the case of datasets earmarked for comparative research.

### 4.3 Concept analysis in ELSST

In ELSST, the English language terms were used as the source for all other language equivalents. This was not to impose dominance of English over any of the other languages, but to try to ensure that a designation for the same concept was sought in every target language. We refer to ‘target language equivalents’ rather than translations, since finding target language equivalents was viewed as being more akin to a mapping process than a translation process (see Doerr, 2001). Functional equivalence, rather than strict equivalence was aimed for. Equivalent terms are expected to allow retrieval of data on the same topic across archives and languages. In practice this means that target language terms are acceptable that express partial as well as exact equivalence with the source term.(Partial equivalence in ELSST is defined as cases where source language and target language terms are generally regarded as referring to the same concept, but where one of the terms strictly denotes a slightly broader or narrower concept. More information on different types of term equivalence in ELSST can be found in Balkan et.al., (2010)).Another consequence of this strategy is that all language versions of ELSST strive to be 'equally authentic', in the sense adopted by European Union language policy (see Athanasiou, 2006; Fidrmuc et al., 2007; Luttermann, 2011), i.e. treated as originals and not as translations.

In ELSST concept analysis was carried out where difficulties arose in finding target language equivalents. Except in straightforward cases it was not possible to find a target language equivalent for a term without a source term definition or scope note. To find a definition, the source language team had to examine synonyms and related terms, and, where necessary, review the research material that had been indexed with the term in order to pin down its exact meaning. The term, together with its scope note, was then passed to the target language teams to find equivalent terms. They in turn often had to perform concept analysis on candidate target language terms in order to establish the degree of equivalence with the source term. Cross-team collaboration was frequently required to arrive at a solution. All terms and their scope notes were also peer-reviewed by field experts.

Another case where concept analysis was required in ELSST was where efforts were made to align ELSST with existing lexical resources (see Section 5 below).

Developing multilingual lexical resources has much in common with establishing interoperability between monolingual resources. Tudhope, Koch, & Heery (2006) note that "*Mapping is a key requirement for semantic interoperability in heterogeneous environments. Although schemas, frameworks and tools can help, detailed mapping work at the concept level is necessary, requiring a combination of intellectual work and automated assistance. Significant effort is required for useful results*." The examples described below were based on intellectual work alone. As with Tudhope et al., mapping between terms was performed at the concept level.

## 5 Multilingual lexical resources and ELSST

### 5.1 Classifications

International classifications were consulted in the development of ELSST particularly for cases where there was a wide discrepancy in terms due to differences in institutional structures and systems. An example is the domain of education, where the International Standard Classification of Education (ISCED-97) (UNESCO, 2006) was consulted.

#### *5.1.1 The International Standard Classification of Education (ISCED-97)*

International Standard Classifications (ISC), produced by United Nations Educational, Scientific and Cultural Organization (UNESCO), serve as tools for comparisons and are considered to be "[…] *the foundations on which all statistical systems, national as well as international, are built*" (Eurostat, 1999:5). It is intended that comparability in ISC be achieved not only via their structure, but also by means of their conceptual transparency.

ISCED was specifically designed "[…] *to greatly improve the comparability of education statistics – as data collected under this framework will allow for the comparison of educational programmes with similar levels of educational content – and to better reflect complex educational pathways in the OECD indicators*" (OECD, 1999:3). ISCED-97 is used in cross-national surveys for the purposes of statistical reporting, creation of research tools, comparisons, policy making, etc. It has been extensively tested for its cross-national applicability (see for example Eurostat, 1999; OECD, 1999; Schneider, 2008). National statistical institutes (NSI) provide resources on educational subjects in compliance with the ISCED-97 levels of education (Eurydice, 2012).

Mapping the ELSST EDUCATION hierarchy, which includes EDUCATIONAL SYSTEMS and EDUCATIONAL INSTITUTIONS, to ISCED-97 demanded long discussions among source and target language teams about educational concepts and their definitions. Specialised resources, such as the European Glossary on Education were also consulted. In view of the difficulties in mapping ELSST to ISCED-97, restructuring ELSST according to ISCED-97 levels was discussed but rejected, on the grounds that little would be gained.

Most problems arose from the fact that ISCED-97 refers to systems that are too culture-specific. In ELSST it was not always possible to find equivalent terms in one or more of the target languages since the corresponding concept was missing. Several solutions were adopted in such cases, such as adding metadata support (i.e.: definitions and scope notes) or employing descriptive phrases or neologisms[4]. For example, consider the term SECONDARY SCHOOLS and its scope note in ELSST (Figure 3).

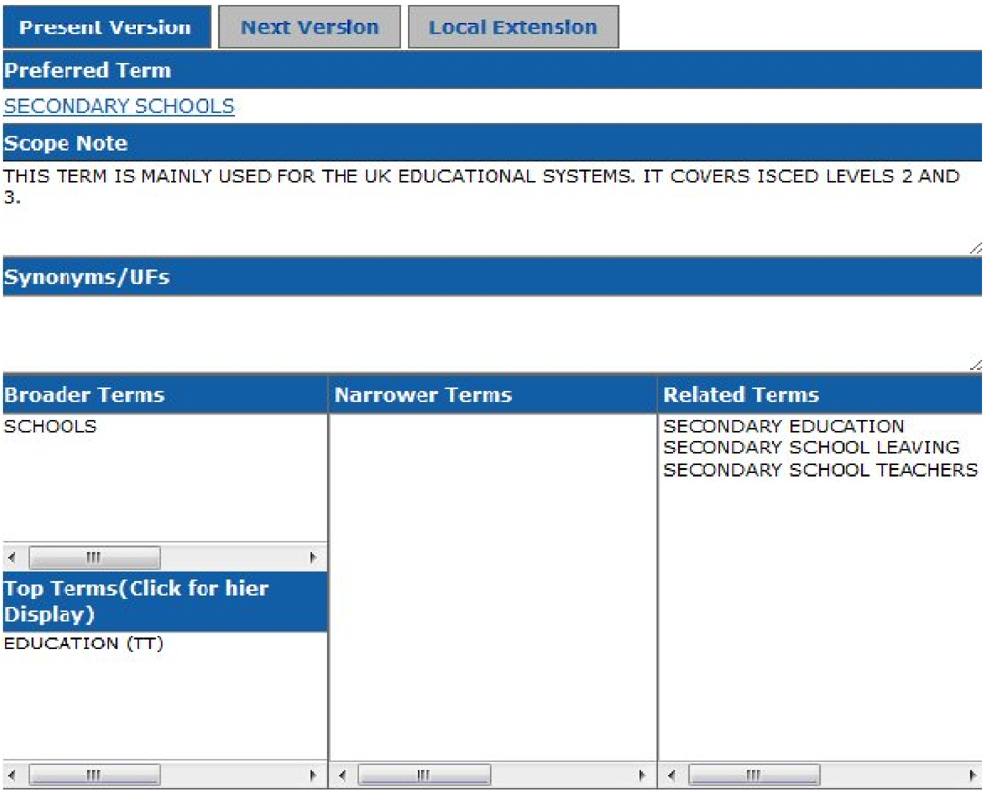



**Figure 3:** ELSST term 'SECONDARY SCHOOLS' in source language

ISCED distinguishes between seven levels of education ranging from pre-primary to tertiary. Secondary education covers ages 11 or 12 through 18 or 19 and is divided into two levels: lower and upper secondary (levels 2 and 3). The English term maps quite well to ISCED

[4] Unlike terminologies, which are only interested in real language use, multilingual thesauri may use synthetic terms or descriptive phrases to bridge lexical gaps, since they provide a reference point for other related terms that can be used to aid search and browsing, if not indexing.

levels 2 and 3, as shown in the scope note of the English term (Figure 3). However, in contrast to English, the educational systems for many target languages are structured differently with respect to these levels. For example, when mapping the Greek school system to ISCED educational levels a problem arises since the term ‘secondary’ is only used in relation to education levels in Greek, not schools (see Figure 4). Thus in ELSST, the synthetic term “ΣΧΟΛΕΙΑ ΔΕΥΤΕΡΟΒΑΘΜΙΑΣ ΕΚΠΑΙΔΕΥΣΗΣ” (literally, schools for secondary level education) is used as the Greek equivalent of SECONDARY SCHOOLS, with “ΓΥΜΝΑΣΙΑ” (corresponding to secondary schools of ISCED level 2) and “ΛΥΚΕΙΑ” (corresponding to secondary schools of ISCED level 3) as its synonyms.

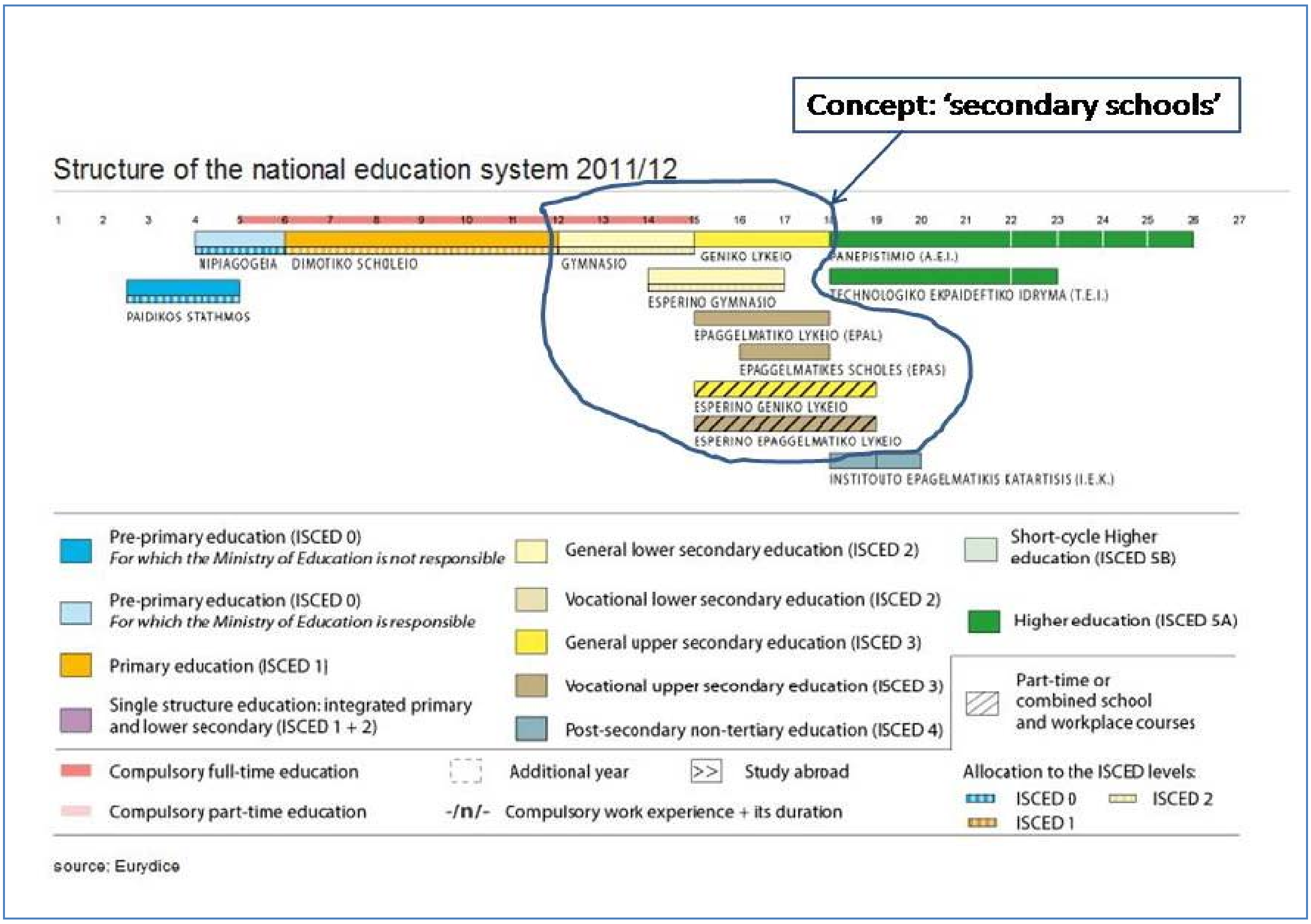



**Figure 4.** An example of mapping ISCED Levels 2 & 3 to the concept ‘Secondary Schools’

Besides the conceptual and linguistic difficulties in mapping ELSST with ISCED another problem lay in the scope for which ISCED was constructed: “*The ISCED was not specifically developed for social survey research and facilitating comparative sociological investigations, but rather for statistical reporting, monitoring and information on national educational policy”* (Schneider, 2008:316). ELSST on the other hand was designed for research and consequently uses research terminology.

Other problems with ISCED noted by researchers (e.g. Ganzeboom, 2008) include an insufficient number of categories, a very slow updating schedule, and the fact that it is based on too small a number of educational systems.

### 5.2 Terminologies

Multilingual term banks were consulted when looking for term definitions and/or multilingual term equivalents in ELSST. An example is EURODICAUTOM, the term bank of the European Commission which is now integrated into the Interactive Terminology for Europe (IATE) (European Union, 2012b) term bank.

#### *5.2.1 EURODICAUTOM*

EURODICAUTOM was the multilingual term bank of the European Commission's translation service, covering all EU language pairs, although not all terms were available in all languages. It covered many different subject domains encompassing the fields of activity of the European Union. Terms were often, but not necessarily, provided with a definition. While this often helped to clarify the term in the source language, it was not always enough to help find a functionally equivalent term in the target language.

An example of where EURODICAUTOM was not helpful in finding a target language equivalent is the English term INNER CITIES. In ELSST, the English term has the scope note: "*The central area of a city, especially if dilapidated or characterised by overcrowding, poverty, etc. (OED)*"

This social phenomenon does not transfer to France, where it is the suburbs, rather than the inner cities which tend to suffer social deprivation. EURODICAUTOM gave a range of French equivalents for the English term INNER CITY, including the literal CENTRE-VILLE, without capturing the connotation difference between the terms in the two languages. The solution adopted in ELSST is to leave the English term un-translated in the French version of ELSST, but to provide it with French synonyms (QUARTIERS DESHERITES, QUARTIERS PAUVRES and VIEUX QUARTIERS – literally DEPRIVED AREAS, POOR PARTS OF TOWN and OLD PARTS OF TOWN) and a translation of the English scope note.

| Preferred Term |
|---|
| INNER CITIES |
| **Scope Note** |
| ZONE CENTRALE D'UNE VILLE, SURTOUT SI ELLE SOUFFRE DE DELABREMENT, SURPEUPLEMENT, PAUVRETE, ETC. |
| **Synonyms/UFs** |
| QUARTIERS DESHERITES<br>QUARTIERS PAUVRES<br>VIEUX QUARTIERS |

**Figure 5.** Entry for French equivalent of INNER CITIES in ELSST

In general, EURODICAUTOM was of limited use, due to the following factors:

- gaps in coverage – many ELSST terms are phrasal, and were not found in EURODICAUTOM
- difficulty in choosing between closely related sub-domains

- lack of clarity concerning term meanings - even where terms had definitions it was not always possible to establish whether they were good equivalents
- lack of information about meaning connotations

On the plus side, the source of terms was frequently provided, which is useful, especially in the case of standardised terminology. The wide variety of near-synonyms in EURODICAUTOM provided useful suggestions for non-preferred terms in ELSST. To compensate for the lack of coverage, automatic term extraction could be investigated (CabréCastellví, EstopàBagot, & Palatresi, 2001).

**5.3 Other thesauri**

Various related thesauri were consulted during the construction of ELSST, including the following: EUROVOC (European Union, 2012a), the UNESCO thesaurus (UNESCO, 2012) and the ILO thesaurus (ILO, 2012) of the European Commission, UNESCO and International Labour Organization (ILO) respectively.

The UNESCO thesaurus covers the domains of education, culture, natural sciences, social and human sciences, communication and information and is used within the organisation and beyond. EUROVOC focuses on the law and legislation of the European Union (EU) and is used in the Library of the European Parliament, the Publication Office as well as other information institutions of the EU. The ILO thesaurus focus is on topics related to economic and social development. All three thesauri are used for indexing and retrieval. ILO is reportedly also used by editors and translators as a specialised glossary.

Some more specialised thesauri were also consulted, including the thesaurus of the Education Resources Information Center (ERIC) (ERIC, 2012), used to index journal articles and other education-related materials, and the European Training Thesaurus (ETT) (Cedefop, 2012), the thesaurus of the European Centre for the Development of Vocational Training (Cedefop)which is used for indexing information relating to vocational education and training systems and programmes, policy, economical and social aspects and training-related EU actions.

While it was useful to consult other thesauri with social science coverage, differences in their scope and conceptual structure meant that it was not always possible to establish (a) whether two terms in the same language referred to the same concept in different thesauri and hence (b) whether the target language equivalent for a term in another thesaurus was suitable for the same term in ELSST. Sometimes it was only by examining the target language equivalent for a term that the difference in meaning between the source terms in two different thesauri emerged.

Term meanings in thesauri are conveyed in a number of ways, including indirectly, via their place in the hierarchy, their synonyms or related terms, or directly, via their scope notes or definitions. Disregarding other factors, Table 2 below illustrates how the terms EDUCATION, TEACHING METHODS, EDUCATIONAL SCIENCES and INSTRUCTION in six different thesauri have some overlap of meaning, based on their synonyms, especially the synonym PEDAGOGY.

| **ELSST** | **EUROVOC thesaurus** | **UNESCO thesaurus** | **European Training Thesaurus** | **ERIC thesaurus** |
|---|---|---|---|---|
| EDUCATION<br>UF:<br>EDUCATIONAL SCIENCES,<br>**PEDAGOGY,**<br>SCHOOLING | EDUCATION | EDUCATION | EDUCATION | EDUCATION |
| TEACHING METHODS | TEACHING METHOD<br>UF =<br>**PEDAGOGY** | TEACHING METHODS<br>UF:<br>EDUCATIONAL METHODS,<br>INSTRUCTIONAL METHODS,<br>TEACHING STRATEGIES,<br>TEACHING TECHNIQUES | (UF of TRAINING METHOD) | - |
| - | - | EDUCATIONAL SCIENCES<br>UF: **PEDAGOGY** | SCIENCES OF EDUCATION<br>UF = **PEDAGOGY** | - |
| (UF of TEACHING) | - | (UF of TEACHING) | - | INSTRUCTION<br>UF: **PEDAGOGY** |

**Table 2.** Partial equivalence of terms between thesauri based on UFs

Further investigation of these terms (by, for example, comparing their scope notes, if any, and their place in the hierarchy) was necessary to establish the degree of equivalence of the concepts expressed by these terms in each thesaurus.

An example of an ambiguity in the source term that was not apparent at a monolingual level is the English term DRUG(S) which is shown with its French equivalents in three separate thesauri in Table 3.

| **ELSST** | **ILO thesaurus** | **UNESCO thesaurus** |
|---|---|---|
| DRUGS<br>FR: DROGUES ET MEDICAMENTS | DRUG<br>FR = DROGUE | DRUGS<br>PHARMACEUTICALS<br>FR = MÉDICAMENT |

**Table 3.** Partial equivalence of terms between thesauri based on their translations

Superficially it looks as though the English term DRUG(S) is referring to the same concept in all three thesauri. However, inspection of the French language equivalents shows that this is not the case. 'MÉDICAMENT' (literally medicinal drugs) is the French equivalent for the term in the UNESCO thesaurus, while DROGUE (literally illegal drug) is the French equivalent in the ILO thesaurus. In ELSST, DRUGS can refer to both legal and illegal drugs, hence the decision to use DROGUES ET MEDICAMENTS (literally ILLEGAL DRUGS and MEDICINAL DRUGS) as its French equivalent.

In summary, interoperability between thesauri is limited by the following:

1. differences in scope and coverage
2. differences in conceptual framework

3. lack of term definitions
4. lack of clarity of meaning of term relationships

Problem (3) can be addressed by providing scope notes and definitions. This is particularly important in a multilingual context, and the first step in finding target language equivalents in ELSST was often to precisely define the meaning of the term in English. Problem (4) is a characteristic of many thesauri, where, for example, the BT relationship can express a number of semantic relationships, not strictly the 'isa' relationship where all objects that belong to the extension of the NT also belong to the extension of the BT. (For example, in ELSST, EDUCATIONAL PERSONNEL is the BT of TEACHERS, where it is the case that all teachers are also educational personnel. However, CHILDHOOD is the BT of CHILDREN, although children are not the same type of thing as childhood.) RT, similarly, can express a number of relationships, including cause and effect (e.g. OFFENCES in ELSST is related to DELINQUENCY and PUNISHMENT). A possible 'solution' is to convert thesauri to ontologies, where relationships between terms are more formally defined, and this is the subject of ongoing research (see for example Hahn (2003) and Biasiotti & Fernandez-Barrera(2009)).

## 6 Conclusion

To return to our original question about the value of re-using lexical resources in the construction of ELSST, our answer is generally affirmative, while acknowledging the difficulty of the task. Given the amount of time and effort that has gone into creating these resources in the first place, it makes sense to re-use them where possible, particularly in the case of standardised terminology. This also promotes semantic interoperability between information sources.

On the negative side, it involves a lot of work, and we encountered a variety of problems in mapping terms (via their concepts) and providing target language equivalents, due to, among other things, the differences in scope and subject coverage between lexical resources employed, the differences in conceptual framework, and the lack of definitions which are of great importance for social sciences terms.

Specifically, we conclude that the following features are desirable in any lexical resource in order to facilitate generation of target language terms, and promote its re-usability:

1. Terms must be clearly defined

Term definitions or scope notes, although not imperative in monolingual resources, are essential in multilingual resources, otherwise mapping errors can arise. Definitions also lead to a better (and more re-usable) end-product. However, although scope notes facilitate concept mapping, they are not adequate in all cases –that is, concept analysis is sometimes required. This raises doubts about the feasibility of fully automating the process of aligning lexical resources, an area of current research.

2. Terms must be unambiguous

This is a general desideratum for thesaurus construction, though not always achievable. As discussed, the process of finding a target equivalent for a term often sheds light on the nature of the underlying concept that was not apparent in the source language, and can thus help clarify the concept. The greater number of target languages, therefore, the greater, potentially, is the clarity of the concepts.

ANSI/NISO Z39.19-2005, (2005) notes that these two conditions are key to interoperability in controlled vocabularies. However, we found that even given these two conditions, exact equivalence between terms is not always possible. This is particularly evident in multilingual thesauri, where there may be cultural differences between concepts.

**Acknowledgements**
We would like to thank all CESSDA archive colleagues who have participated in the construction of ELSST.

**7 References**

Alvheim, A. (2006): *Multilingual access to data infrastructres of the European Research Area*. Research Report, CESSDA, Norwegian Social Science Data Services: Bergen, Norway.
ANSI/NISO Z39.19-2005 (2005):*Guidelines for the construction, format, and management of monolingual controlled vocabularies.*NISO Press: Bethesda, Maryland, U.S.A.
Athanasiou P. (2006): The Application of Multilingualism in the European Union Context. *Legal Working Paper Series,* No.2, (March). European Central Bank. http://www.ecb.int.
Balkan, L. et al. (2010): European Language Social Science Thesaurus (ELSST): Issues in designing a multilingual tool for social science researchers. *COST A31 Final Conference: Categorising Human Experience: Classification in Languages and Knowledge Systems.* Paris.
Biasiotti, M. A. & Fernandez-Barrera, M. (2009): Enriching thesauri with ontological information: Eurovoc Thesaurus and DALOS Domain Ontology of Consumer Law. *Workshop on Legal Ontologies and Artificial Intelligence Techniques, joint with 2nd Workshop on Semantic Processing of Legal Texts. In Conjunction with ICALL 2009, 12th International Conference on Artificial Intelligence and Law.* Barcelona8-12 June. http://idt.uab.es/images/IDT_Collections/IDT_Series/IDTSeries2_LOAIT.pdf.
CabréCastellví, M. T., EstopàBagot, R.& Palatresi, J. V. (2001): Automatic term detection: a review of current systems. In *Recent Advances in Computational Terminology* (pp. 53-87). John Benjamins Publishing Company: Philadelphia, PA.
Campbell, K. E., & Giannangelo, K. (2007): Language barrier: getting past the classifications and terminologies roadblock. *Journal of AHIMA,* 78(2).
Cedefop (2012): *European Training Thesaurus (ETT)* European Centre for the Development of Vocational Training. http://libserver.cedefop.europa.eu/ett/en/
CESSDA PPP. (2008-2010): *WP4 Synopsis.* http://www.cessda.org/project/doc/wp04synopsis_sept08.pdf
Chouliaraki, L. & Fairclough, N. (1999):*Discourse in late modernity – rethinking critical discourse analysis.* Edinburgh University Press: Edinburgh.
DDC. (1989):*Dewey Decimal Classification -DDC20.* Forest Press: New York.
DDI Alliance. (2009):*What is DDI?*http://www.ddialliance.org/what
Doerr, M. (2001): Semantic problems of thesaurus mapping. *Journal of Digital Information,* 1(8).
ERIC (2012): *ERIC thesaurus.* European Resources Information Centre
http://www.eric.ed.gov/ERICWebPortal/resources/html/thesaurus/about_thesaurus.html
European Union (2012a): *Eurovoc thesaurus.* http://eurovoc.europa.eu/
European Union (2012b): *Interactive Terminology for Europe* (IATE) term bank. http://iate.europa.eu/

Eurostat (1999): *Inventory of International Statistical Classifications* (1999 ed., Vol. Theme 2: Economy & Finance). European Communities: Luxembourg.
Eurydice (2012): *European Glossary on Education* (2nd ed., 2: Educational Institutions).
Eurydice, the Information Network on Education in Europe: Brussels.
Fidrmuc, J., Ginsburgh, V. and Weber, S. (2007): Ever Closer Union or Babylonian Discord? The Official-language Problem in the European Union, *Discussion paper*, Centre for Economic Policy Research. http://www.fidrmuc.net/research/FGW2.pdf
Ganzeboom, H. B. (2008): Harmonising education and occupation in cross-national comparative research. *CESSDA PPP Expert Workshop on Harmonisation Issues in Comparative Social Surveys*. Paris, France.
Granda, P., Kramer, S.& Linnerud, J. (2008): Controlled vocabulary. In K. M. Stefan Kramer (ed.) DDI Alliance.
Guarino, N. (1995): Formal ontology, conceptual analysis and knowledge representation. *International Journal of Human Computer Studies,* 43 (5-6): 625-640.
Guédon, J. C. (2008): Open access: restoring the Republic of Science and its Great Conversation. *International Conference : Open Access Infrastructures: The Future of Scientific Communication.* Athens: National Hellenic Research Foundation.
Hahn, U. (2003): Turning informal thesauri into formal ontologies: a feasibility study on biomedical knowledge re-use, comparative and functional genomics. *Comp Funct Genom,* 4: 94-97*(DOI: 10.1002/cfg.247).*
Harbsmeier, C. (2007): *Concepts that make multiple modernities: the conceptual modernisation of China in a historical and critical perspective.* http://www.hf.uio.no/ikos/english/research/projects/tls/publications/CONCEPTS-THAT- MAKE-HISTORY%5B1%5D.pdf
Hayek, F. A. (1976): *The sensory order: an inquiry into the foundations of theoretical psychology*. The University of Chicago Press: Chicago, IL.
Hodge, G. (2000): *Systems of knowledge organisation for digital libraries: beyond traditional authority files.* The Digital Library Federation Council on Library and Information Resources: Washington DC.
ILO (2011): *International Standard Classification of Occupations (ISCO).*International Labour Organization http://www.ilo.org/public/english/bureau/stat/isco/isco08/index.htm
ILO (2012): *ILO thesaurus*. International Labour Organizationhttp://www.ilo.org/thesaurus
ISO (1985*): Documentation guidelines for the establishment and development of multilingual thesauri.International Standard -ISO(5964).* International Standardisation Organisation.
ISO (1986): *Documentation Guidelines for the Establishment and Development of Monolingual Thesauri.International Standard -ISO(2788)*. International Standardisation Organisation: Geneva.
ISO (1990): *Terminology - vocabulary ISO (1087).* International Standardisation Organisation: Geneva.
ISO (2009): *ISO 704:2009 Terminology work - principles and methods*. International Standarsisation Organisation: Geneva. ftp://ftp.omg.org/pub/sbvr rtf/ISOStandards/ISO704_2009.pdf
Kappi, C. (2009): Diachrony of scientific classifications: the case of change in social classification systems. *Address at Conference: "The Diachrony of Classification Systems"*. 12-13 March, Wassenaar, NL: COST A-31.
Lauser, B. et al. (2008): Comparing human and automatic thesaurus mapping. *Proc. Int'l Conf. on Dublin Core and Metadata Applications 2008*, http://edoc.hu-berlin.de/conferences/dc-2008/lauser-boris-43/PDF/lauser.pdf.

Luttermann K. (2011): Cultures in Dialogue. Institutional and Individual Challenges for EU Institutions and EU Citizens from the Perspective of Legal Linguistics. *Hermes – Journal of Language and Communication Studies*, No. 46.

Miller, K. (2004): No Longer Lost in Translation. *IASSIST 2004 -Data Futures: Building on 30 Years of Advocacy*. IASSIST: Madison, Wisconsin USA. http://www.iassistdata.org/downloads/2004/e1_miller.pdf

Miller, K.& Matthews, B. (2001): Having the right connections: the LIMBER project. *Journal of Digital Information*, 1(8).

Nuopponen, A. (2010a): Methods of concept analysis - a comparative study. *LSP Journal*, 1(1): 4-12.

Nuopponen, A. (2010b): Methods of concept analysis - towards systematic concept analysis. *LSP Journal*, 1(2):5-14.

Nuopponen, A. (2011): Methods of concept analysis - tools for systematic concept Analysis. *LSP Journal*, *2*(1): 4-15.

OECD (1999): Classifying Educational Programmes*Manual for ISCED-97 Implementation in OECD Countries* Organisation for Economic Co-operation and Development.

Priss, U. (1996): *Relational Concept Analysis: Semantic Structures in Dictionaries and Lexical Databases.* Technischen Universit□at Darmstadt, Fachbereich Gesellschafts- und Geschichtswissenschaften. Darmstadt: http://www.upriss.org.uk/papers/diss.pdf.

Priss, U. (2006): Formal concept analysis in information science. *Annual Review of Information Science and Technology, ASIST*, 40.

Schneider, S. L. (Ed.) (2008): *The International Standard Classification of Education (ISCED-97).*Mannheimer Zentrum für Europäische Sozialforschung: Mannheim, Germany.

Schneider, S. L. (2008): Suggestions for the cross-national measurement of educational attainment: refining the ISCED-97 and improving data collection and coding procedures. In S. L. Schneider (ed.), *The International Standard Classification (ISCED-97). An evaluation of content and criterion validity for 15 European countries* (pp. 311-330). Mannheimer Zentrum für Europäische Sozialforchung (MZES): Mannheim.

The Library of Congress. (2009, 8 10): *Library of Congress Classification*. http://www.loc.gov/catdir/cpso/lcc.html

Tudhope, D., Koch, T., & Heery, R. (2006): *Terminology services and technology: JISC state of the art review*. Project Report: Joint Information Systems Committee, Opus: University of Bath Online Publication Store. http://opus.bath.ac.uk/23563/1/terminology_services_and_technology_review_sep_06.pdf

UNESCO (2006): *International Standard Classification of Education (ISCED-97).* United Nations Educational, Scientific and Cultural Organization.

UNESCO (2012): *UNESCO thesaurus* United Nations Educational, Scientific and Cultural Organization.http://www2.ulcc.ac.uk/unesco/

University of Essex (2012): *Humanities and Social Science Electronic Thesaurus* (HASSET). UK Data Archive, University of Essex http://www.data-archive.ac.uk/find/hasset-thesaurus

van Hage, W. R. et al. (2008): The OAEI food task: an analysis of a thesaurus alignment task in *Applied Ontology*, 1(1).*IOS Press*http://www.cs.vu.nl/~guus/papers/Hage10d.pdf.

WHO(1998): *The World Health Report 1998 -Life in the 21st century; A Vision for All*, WHO Director General.

WHO (2010a): *International Classification of Diseases (ICD).*World Health Organization. http://www.who.int/classifications/icd/en/

WHO (2010b): *International Classification of Functioning Disability and Health (ICF)*. World Health Organization. http://www.who.int/classifications/icf/en/

Zeng, M. L. & Hodge, G. (2011): Developing a Dublin Core application profile for the Knowledge Organisation Systems (KOS) resources. *ASIS&T Bulletin*, April/May. http://www.asis.org/Bulletin/Apr-11/AprMay11_Zeng_Hodge.html.

***

LSP Journal, Vol.3, No.1 (2012) / http://lsp.cbs.dk

# Analysing and teaching keywords in hotel brochure text

Wenhsien Yang

Department of Applied English,
National Kaohsiung University of Hospitality and Tourism,
Kaohsiung, Taiwan
yangwenhsien@mail.nkuht.edu.tw

***Keywords:*** *analysis of promotional genre, genre-based writing instruction, hotel brochures, keyword analysis*

**Abstract**

This study examined written texts in English from the hospitality industry. Authentic printed hotel brochures together with ESP learners' productions from a genre-based writing instruction course were collected, forming the two corpuses to be examined. A corpus analysis generated wordlists as well as keywords. The results showed that the students' choices of lexis differed to some extent from the word use in the authentic texts. This finding suggests that in a genre writing class instructing ESP learners in how to compose authentic English informative and promotional texts, priority needs to be given to teaching keywords, as these are the words that an English-speaking audience would normally expect to read in such texts. In addition, the genre-based instruction of analysing keywords helps learners transform their personal-style texts into more conventionally-accepted texts. In conclusion, this paper argues that subtle choices and constraints make the information in hotel brochures authentic, persuasive, trustworthy, and motivating.

## 1 Introduction

Recently, a great number of ESP-related studies on professional disciplines such as medical English, legal English, commercial English, and academic English have been explored in the context of applied linguistics (Bhatia, 1993; Hyland, 2004b; Swales, 1990; Vergaro, 2004). English for hospitality and tourism (E*f*H&T), including catering, tourism and hotel management, has also started receiving equal attention from researchers (Henry & Roseberry, 1996, 1998; Moya Guijarro & Albentosa Hernandez, 2001; Lam, 2006; Moya Guijarro, 2006; Lien, 2008).

Based on the policies of national development, the hospitality and tourism (H&T) industries are key areas which are to be highly promoted within the next few years by Taiwanese authorities. Prospective professionals in these two areas are expected to be equipped with proficiency in English for international communication, in accordance with the current globalisation phenomenon. Enhancing students' English proficiency for international communication or for social purposes has thus become the top priority for English educators in poly-technical universities.

It has always been assumed that English for hospitality and tourism (E*f*H&T) should focus on listening and speaking skills, as these are the two skills that students are most likely to need in their future workplaces. Oral communication skills are indeed very practical in the workplace, especially for those who have to serve their guests face to face. Therefore, most of the published textbooks about English in this professional area are oriented toward oral communication, with few focusing on reading, not to mention writing skills.

English writing for hotel brochures is not currently a required skill for employees, and may be sacrificed by educational institutes due to the rapid pace of change in these industries, which makes it difficult to keep up with the latest requirements. However, this paper argues that the importance of writing skills in the H&T industry should not be underestimated because writing is permanent and it is expensive to correct written errors or mistakes in comparison with oral mistakes (Wildes & Nyheim, 2009). Besides, written information is often the initial source of information for guests, and creates the first impression of the services provided, even before the guests arrive at the venue. Accordingly, how to express hospitality accurately in written form makes English writing in the hospitality industry an important skill.

In this study, the focus is placed on analysing written English in hotel brochures, a specific genre used as an example to explore the basic component – lexis in teaching genre writing. Words are one of the most essential elements in writing, and various professions tend to emphasise the use of different keywords, making their written texts look authentic and fitting for the specific context. It is believed that a focused examination of lexical-grammatical features in a language can facilitate ESL/EFL learners' acquisition of appropriate vocabulary, as it allows them to compare authentic choices with their inter-language hypothesis (Ellis, 1995; Wennerstrom, 2003). Thus, it is assumed that priority should be given to the analysis and learning of the particular keywords of a genre if learners are expected to write professionally for specific purposes. In this study, the frequently-used words in authentic hotel brochures were compared with lists of keywords generated from students' writing. The goal of this project was to evaluate whether or not the students would perform better in their brochure writing after a teacher explicitly analysed the keywords, including the overused and underused words, in a genre-based instructional (GBI) context.

## 2 Literature Review

In lexicogrammar analysis, the concepts of register and genre are often relied on; however, the differentiation of these two terms is not always easy. Thus, they are sometimes used interchangeably because they overlap to some extent (Lee, 2001; Biber & Cornard, 2009). There is specific relationship between genre and register and they are sometimes complementary to each other. As Ventola (1995: 7) argues, "system choices from the genre networks are realised by the choices on the register, which in turn are realised by the language plane." According to Crystal (1191: 295), register is defined as 'a variety of language defined according to its use in social situations, e.g. a register of scientific, religious, formal English," and it closely associates with the configuration of tenor, mode and field in Hallidaian grammar (Halliday, McIntosh, & Strevens, 1964). Hence, register analysis involves examining situations, linguistic features and functions of texts. However, genre is generally considered to be a level higher than register (Martin, 1993). According to Swales (1990), a genre consists of a series of goal-oriented communicative events formed out of schematic structures, and members in these events share the same communicative purposes. A genre is highly structured and conventionalised, and has specific constraints such as lexis and moves exploited by the members in a specific community to achieve communicative purposes

(Bhatia, 1993). In contrast to register, genre tends to be connected with the organisation of culture and social purposes around language (Bhatia, 1993; Swales, 1990). Corpus studies on conventionally recognisable texts built upon the concepts of genre can better attend to the dynamic/negotiated aspects of situated language use (Lee, 2001). Hence, analysing a genre can help ESP practitioners and learners identify how texts are structured and distinguished in social and cultural contexts in order to realise their communication purposes (Hyland, 2004a). In addition, genre-based pedagogy offers students "explicit and systematic explanations of the ways language functions in social contexts" (Hyland, 2004a: 18), which helps students acquire the specialist culture (Bhatia, 1997). Both Paltridge (2001) and Hyland (2007) advocate explicit instruction in analysing genres at both a broad (i.e., contextual) and a narrow (i.e., linguistic) level, and providing learners with the knowledge and skills necessary to communicate in particular situations and to gain access to socially powerful forms of language. Research on genre analysis has been incorporated in teaching English for academic purposes (e.g., Carstens, 2008; Cheng, 2006; Henry & Roseberry, 2007; Hyland, 2010; Hyon, 2008; Johns, 1995, 1997, 2007; Maingueneau, 2002; Samraj & Monk, 2008; Wennerstrom, 2003; Woodward-Kron, 2005; Wrigglesworth & McKeever, 2010; Zhu, 2005) and for non-academic (or occupational) purposes (e.g., Hafner, 2010; Henry & Roseberry, 1996, 1998; Karlsson, 2009; Matt, 2007; Osman, 2008; Schneider & Andre, 2005; Wennerstrom, 2003). In particular, teaching genres has been shown to have a positive effect on the English writing of L1 Chinese students (Lee & Chen, 2009).

Genre analysis of promotional tourism and hospitality texts has focused on both micro and macro levels. Some studies have explored the issue at the micro level, that is, investigating the linguistic features of travel texts (e.g., Lam, 2006; Moya Guijarro, 2006; Moya Guijarro & Albentosa Hernandez, 2001), while other studies have examined texts at both micro and macro levels, that is, combining linguistic features and generic moves (e.g., Henry & Roseberry, 1996; Lien, 2008). In a genre analysis of promotional travel texts, Henry and Roseberry (1996) found four obligatory moves: the Identification, Facilities/Activities, Location, and Description moves. In a comparable study of promotional travel texts in Taiwan, Lien (2008) outlined three moves: Attracting reader's attention, Detailing the product, and Stating information. The move analysis proposed by Lien overlaps with Henry and Roseberry's analysis of moves. For instance, the move *attracting reader's attention* is identical to the *facilities/activities* move, as they both present unique facilities or activities to attract readers. Move analysis provides an overarching framework for lexical analysis of tourist texts.

In order to write persuasive promotional travel texts, lexis is an important component. For example, subjective language such as emotional, commentary or axiological adjectives, is commonly used, and the consistent use of imperatives and simple present tense are also major features (Lam, 2006; Lien 2008; Moya Guijarro, 2006). To express hospitality and provide a sense of inclusion, first person pronouns such as *we, our* and *I* are usually seen in tourism texts (Lien, 2008; Yang, 2010a, 2010b). Surprisingly, common verbs such as *be, take, find* and *enjoy* also appear frequently. In addition to these linguistic features, Lien (2008) and Moya Guijarro and Albentosa Hernandez (2001) emphasise the importance of including iconic elements in travel texts in order to motivate readers. However, only Henry and Roseberry (1988) have conducted an evaluation of a genre-based approach on the teaching of tourist information in a Brunei university, from which they concluded that the genre group improved significantly on their motivation and texture scores compared to the non-genre group.

In contrast to most of the previous research on examining travel texts, the present experimental study focuses on the genre of hotel brochure text. This genre has received comparatively little attention in previous research projects; therefore, this study combines analysis of the genre with an investigation of the effectiveness of explicit genre-based instruction in an ESP course. Lexis, and keywords in particular, is believed to play a central role in authorship attribution and statistical approaches to style (Davis, 2009; Kirk, 2009). Keywords are important because they reflect what the text is really about in a targeted situation, "avoiding trivia and insignificant detail" (Scott & Tribble, 2006: 56). Hence, teaching keywords in a genre helps learners identify differences between texts, determine the content of texts, and identify textual and rhetorical styles (Archer, Culpeper, & Rayson, 2009; Baker, 2009).

## 3 Methodology

In 2010, 50 authentic hotel brochures written in English were collected from the UK, mainly from tourist bureaus, hotels and information centres in London and Birmingham. All the brochures were gathered by the researcher by visiting authentic contexts/sites in person to ensure the data gathered were the latest and were currently being dispensed to visitors. Most of the brochures were printed double-sided on a single page folded in three.

The students' texts all came from the assignments of 24 English majors taking an 18-week genre-based writing course called *English Writing for Hospitality and Tourism* at a national university in Taiwan. All of the students, who had gained TOEIC scores of 750 to 980, had completed a one-year job placement in either a domestic or foreign H&T company, and this was their final year before graduation. From September 2010 to January 2011, the course was delivered using genre-based instruction (informative and promotional brochures, I&P*b*) through analysing the genre at both macro and micro levels, as suggested by Paltridge (2001). During this three-part course, the students were required to submit three different H&T brochures, that is, hotel brochure #1 (the 1st phase), hotel brochure #2 (the 2nd phase) and a holiday brochure (the 3rd phase), for assessment purposes at the end of each phase. Each brochure was revised and resubmitted three times (termed as V1, V2, and V3) based on the instructor's feedback. The data collected for this research project came from the second phase (i.e., hotel brochure #2). Therefore, there were 13 texts for the first version and 24 texts for the last two versions, giving 61 texts in total[1]. In order to identify the changes, each version was analysed and compared to the authentic texts in terms of lexicogrammar usage at the micro level of genre instruction. Table 1 summarises the structure of the course.

After the texts were collected, all the words in both the authentic and the students' brochures were transcribed and saved in text format in order to be analysed by the computerised tool, WordSmith v.5.0 (WS). Then, wordlists were created. The two sets of texts, that is, the 50 authentic texts and the 24 students' texts, were imported to the WS working area and calculated separately. WS generated two wordlists, one for the authentic brochures and the other for the students' brochures. The lists can be displayed either in frequency order or alphabetical order, showing how many times a word occurs in all texts and the percentage of its appearance. Figure 1 is a sample of the wordlist of the authentic brochures ranked by frequency, while Figure 2 is the students' V2 wordlist.

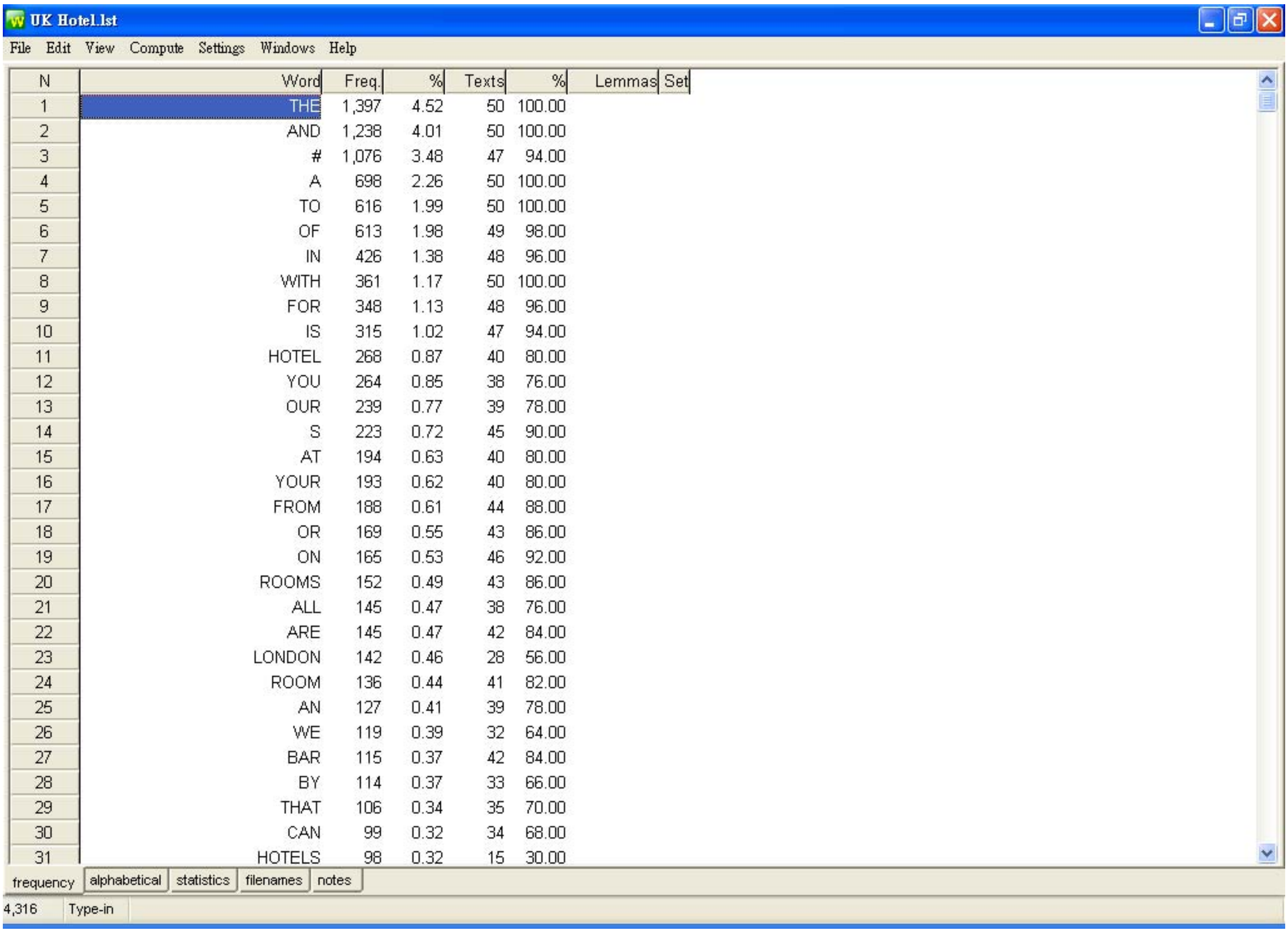
[1] Only 13 students submitted the 1st version of the texts but all 24 students submitted both the 2nd and the 3rd versions of the texts after they were compulsorily required by the instructor.

Next, the keywording process was carried out, following a similar procedure to wordlisting. The major difference lies in the fact that there should be two wordlists in order to generate a keyword list. One is the study/main corpus from which keywords could be identified, and the other is a wordlist serving as the reference/comparison corpus. The word size of the latter corpus is usually much larger than that of the first one and also more general in terms of word selection. The reason why there are two different corpora is because keywording is used to examine items with unusual frequencies (Scott & Tribble, 2006), which helps researchers locate the specific words in one genre in comparison with a reference corpus. In other words, it is these keywords that characterise a specific language use. In this study, the keywords used for designing an English hotel brochure and students' unusual lexical choices were identified, as they are pedagogically useful in teaching students writing English for H&T purposes.

To generate the keywords of the authentic hotel texts, the British National Corpus (BNC) was used as the reference corpus, and the wordlist of the authentic texts was used for the study corpus. Similarly, to generate the students' overused and underused keywords, the main corpus was a wordlist of their texts and the comparison corpus was the wordlist of the authentic texts. In total, 602 keywords were identified in the hotel texts with reference to BNC, and the top 25 content and functional words are categorised as seen in Appendix 1. Appendix 2 displays all three versions (V1, V2, & V3) of keywords in the students' texts with reference to the wordlist of authentic texts. All shaded words are underused keywords compared to the reference corpus.

| Steps | Descriptions |
|---|---|
| 1. | Authentic texts (AT) collected |
| 2. | The referred corpus established (RC) |
| 3. | Analysing the genre texts at both macro and micro levels |
| 4. | Students' writing the first version of the texts (V1) |
| 5. | The studied corpus established (SC) |
| 6. | Wordlists and keywords identified by comparing RC, SC & BNC |
| 7. | Teachers' explaining contextual differences and teaching keywords |
| 8. | Student's revising (as the second version, V2) |
| 9. | (Repeat steps 5, 6 & 7) |
| 10. | Student's finalising (as the third version, V3) |
| 11. | (Repeat steps 5, 6 & 7) |
| 12. | Teacher's evaluation of generic moves and linguistic choices |
| 13. | English-speaking readers' evaluation of motivation and authenticity |

**Table 1.** Summary of the course structure



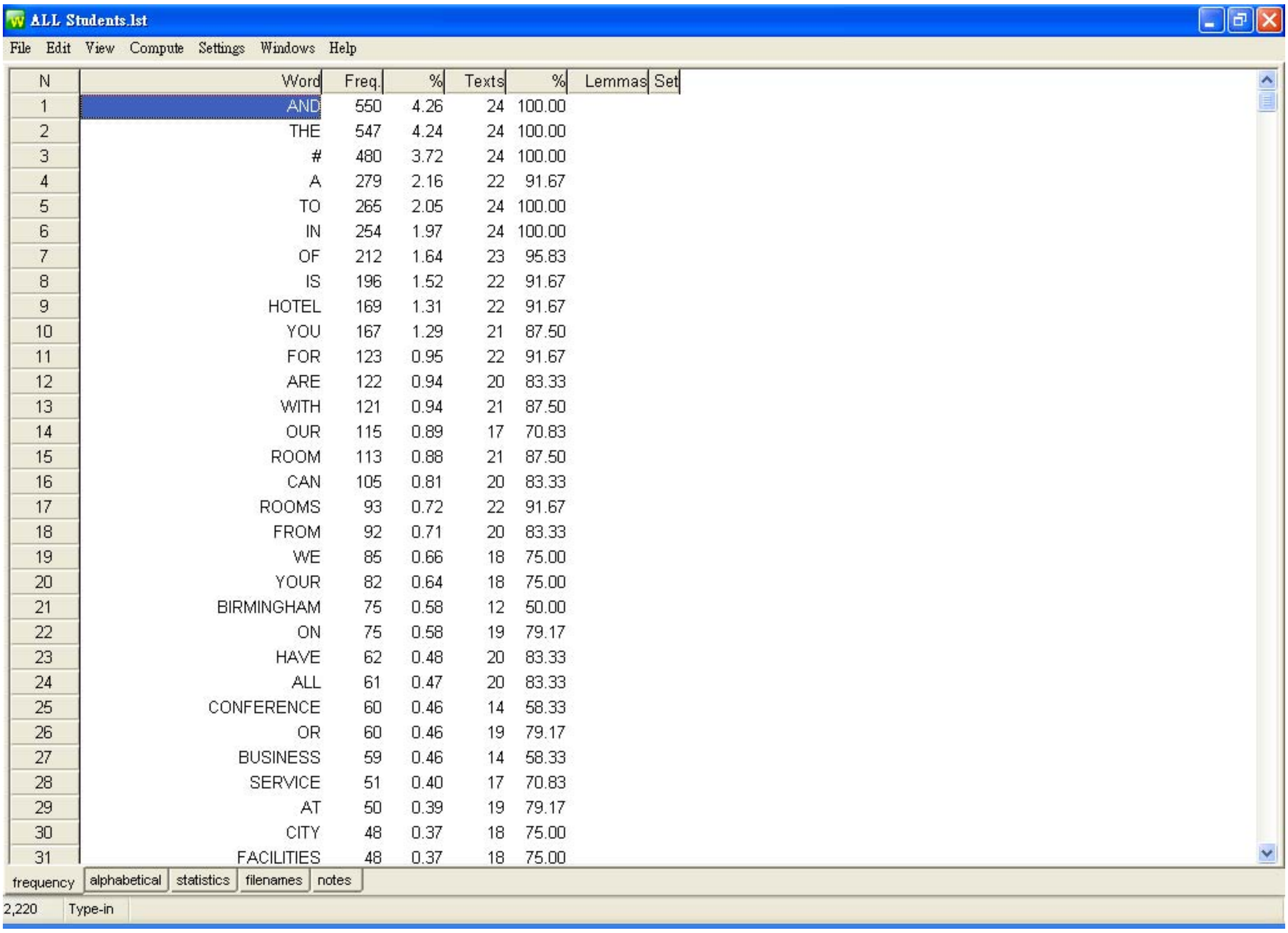

| N | Word | Freq. | % | Texts | % | Lemmas | Set |
|---|---|---|---|---|---|---|---|
| 1 | THE | 1,397 | 4.52 | 50 | 100.00 | | |
| 2 | AND | 1,238 | 4.01 | 50 | 100.00 | | |
| 3 | # | 1,076 | 3.48 | 47 | 94.00 | | |
| 4 | A | 698 | 2.26 | 50 | 100.00 | | |
| 5 | TO | 616 | 1.99 | 50 | 100.00 | | |
| 6 | OF | 613 | 1.98 | 49 | 98.00 | | |
| 7 | IN | 426 | 1.38 | 48 | 96.00 | | |
| 8 | WITH | 361 | 1.17 | 50 | 100.00 | | |
| 9 | FOR | 348 | 1.13 | 48 | 96.00 | | |
| 10 | IS | 315 | 1.02 | 47 | 94.00 | | |
| 11 | HOTEL | 268 | 0.87 | 40 | 80.00 | | |
| 12 | YOU | 264 | 0.85 | 38 | 76.00 | | |
| 13 | OUR | 239 | 0.77 | 39 | 78.00 | | |
| 14 | S | 223 | 0.72 | 45 | 90.00 | | |
| 15 | AT | 194 | 0.63 | 40 | 80.00 | | |
| 16 | YOUR | 193 | 0.62 | 40 | 80.00 | | |
| 17 | FROM | 188 | 0.61 | 44 | 88.00 | | |
| 18 | OR | 169 | 0.55 | 43 | 86.00 | | |
| 19 | ON | 165 | 0.53 | 46 | 92.00 | | |
| 20 | ROOMS | 152 | 0.49 | 43 | 86.00 | | |
| 21 | ALL | 145 | 0.47 | 38 | 76.00 | | |
| 22 | ARE | 145 | 0.47 | 42 | 84.00 | | |
| 23 | LONDON | 142 | 0.46 | 28 | 56.00 | | |
| 24 | ROOM | 136 | 0.44 | 41 | 82.00 | | |
| 25 | AN | 127 | 0.41 | 39 | 78.00 | | |
| 26 | WE | 119 | 0.39 | 32 | 64.00 | | |
| 27 | BAR | 115 | 0.37 | 42 | 84.00 | | |
| 28 | BY | 114 | 0.37 | 33 | 66.00 | | |
| 29 | THAT | 106 | 0.34 | 35 | 70.00 | | |
| 30 | CAN | 99 | 0.32 | 34 | 68.00 | | |
| 31 | HOTELS | 98 | 0.32 | 15 | 30.00 | | |



**Figure 1.** Screenshot of the wordlist of the authentic brochures

The two different keyword lists have different implications. On one hand, the keywords generated from the authentic texts as listed in Appendix 1 are essential words for designing hotel texts for informational and promotional purposes, and these words should be given priority treatment in class. On the other hand, the overused keywords generated from the students' texts should be avoided as they appear too frequently compared with their use in the authentic texts, while the underused words in their texts should be used more often for the opposite reason.

The above are the procedures applied for treating the three versions of the students' texts. Hence, in the end, three versions of keywords were created in order to probe if the students chose lexis differently after the GBI. The method of evaluating whether there were improvements in the texts was to have each student ask an English-speaking reader, mainly native English speakers, to read their three versions to decide which one was better in terms of its authenticity. The following section displays the results and offers likely explanations.




| N | Word | Freq. | % | Texts | % | Lemmas | Set |
|---|---|---|---|---|---|---|---|
| 1 | AND | 550 | 4.26 | 24 | 100.00 | | |
| 2 | THE | 547 | 4.24 | 24 | 100.00 | | |
| 3 | # | 480 | 3.72 | 24 | 100.00 | | |
| 4 | A | 279 | 2.16 | 22 | 91.67 | | |
| 5 | TO | 265 | 2.05 | 24 | 100.00 | | |
| 6 | IN | 254 | 1.97 | 24 | 100.00 | | |
| 7 | OF | 212 | 1.64 | 23 | 95.83 | | |
| 8 | IS | 196 | 1.52 | 22 | 91.67 | | |
| 9 | HOTEL | 169 | 1.31 | 22 | 91.67 | | |
| 10 | YOU | 167 | 1.29 | 21 | 87.50 | | |
| 11 | FOR | 123 | 0.95 | 22 | 91.67 | | |
| 12 | ARE | 122 | 0.94 | 20 | 83.33 | | |
| 13 | WITH | 121 | 0.94 | 21 | 87.50 | | |
| 14 | OUR | 115 | 0.89 | 17 | 70.83 | | |
| 15 | ROOM | 113 | 0.88 | 21 | 87.50 | | |
| 16 | CAN | 105 | 0.81 | 20 | 83.33 | | |
| 17 | ROOMS | 93 | 0.72 | 22 | 91.67 | | |
| 18 | FROM | 92 | 0.71 | 20 | 83.33 | | |
| 19 | WE | 85 | 0.66 | 18 | 75.00 | | |
| 20 | YOUR | 82 | 0.64 | 18 | 75.00 | | |
| 21 | BIRMINGHAM | 75 | 0.58 | 12 | 50.00 | | |
| 22 | ON | 75 | 0.58 | 19 | 79.17 | | |
| 23 | HAVE | 62 | 0.48 | 20 | 83.33 | | |
| 24 | ALL | 61 | 0.47 | 20 | 83.33 | | |
| 25 | CONFERENCE | 60 | 0.46 | 14 | 58.33 | | |
| 26 | OR | 60 | 0.46 | 19 | 79.17 | | |
| 27 | BUSINESS | 59 | 0.46 | 14 | 58.33 | | |
| 28 | SERVICE | 51 | 0.40 | 17 | 70.83 | | |
| 29 | AT | 50 | 0.39 | 19 | 79.17 | | |
| 30 | CITY | 48 | 0.37 | 18 | 75.00 | | |
| 31 | FACILITIES | 48 | 0.37 | 18 | 75.00 | | |

**Figure 2.** Screenshot of the wordlist of the students' V2 brochures

## 4 Results and Discussion

### 4.1 The wordlist for authentic hotel brochures

Table 2 shows that there are 29,830 tokens constituting 4,316 distinct words in the wordlist of the authentic texts. The mean number of words of each brochure is 596.6, and 3-letter words are the most commonly used, followed by 4-letter words, 2-letter words and 5-letter words. In comparison, Table 3 exhibits that there are only 2,140 distinct words of 12,014 tokens on average in the students' 24 texts, and the mean number of words is 590 per text across the three versions. Figure 1 and Figure 2 in the preceding sections partially show the statistics of the texts as they simply display part of the high-frequency words used in the texts. The following Table 2 shows that in general the students' texts were similar in length to the authentic texts, but that they used less than half the number of distinct words when compared with the authentic texts (2,140:4,316).

| | |
|---|---|
| Authentic texts (AT) | 50 |
| Total tokens[2] | 29,830 |
| Distinct words (DW) | 4,316 |
| Keywords (KW) with reference to BNC | 602 |
| DW per 100 tokens | 14.46 |
| KW per 100 tokens | 2.01 |
| KW per 100 DW | 13.94 |
| **Avg. tokens in an AT** | **596.6** |

**Table 2.** Total tokens, distinct words (DW) and keywords (KW) in the authentic texts (AT)

| **Students Texts (ST)** | **The 1st version** | **The 2nd version** | **The 3rd version** |
|---|---|---|---|
| Texts | 13 | 24 | 24 |
| Total tokens[3] | 10,109 | 12,431 | 13,504 |
| Distinct words DW | 1,862 | 2,220 | 2,340 |
| DW per 100 tokens | 18.41 | 17.85 | 17.32 |
| Avg. tokens in a ST | 777.61 | 517.95 | 562.66 |
| **Keywords (KW)** ***Ref. to BNC*** | **202** | **266** | **270** |
| KW per 100 tokens | 1.99 | 2.13 | 1.99 |
| and DW | 10.84 | 11.98 | 11.53 |
| **Keywords (KW)** ***Ref. to AT*** | **20** | **16** | **16** |
| KW per 100 tokens | 0.19 | 0.12 | 0.11 |
| and DW | 1.07 | 0.72 | 0.68 |
| **Keywords excl. special terms** | **20** | **16** | **12** |
| KW[4] per 100 tokens | 0.19 | 0.12 | 0.08 |
| KW per 100 DW | 1.07 | 0.72 | 0.51 |

**Table 3.** Total tokens, distinct words (DW) and keywords (KW) in the student texts (ST)

### 4.2 The keywords in the authentic hotel brochures compared to BNC

In the authentic texts, there are 602 keywords in total, 556 of which are overused words, while the 46 words in are underused compared to BNC. The higher ranking a word has, the more keyness it has in the study corpus. Hence, the word, *hotel* has the highest keyness in these hotel brochures. The 557th to the 602nd ranked words are underused when compared to BNC, which suggests that they are not as essential as the overused words from ranking 1 to 556. Perhaps these 46 words such as *like* and *to* are commonly used in BNC but not in hotel brochures. Hence, the pedagogical implication is that teachers should consider the top 556 words as a priority since they constitute the vocabulary which is most often used in authentic brochures.

[2] This is used for word list as the reference corpus.
[3] This is used for word list as the study corpus.
[4] Special terms are eliminated in this list.

The next step was that the above 602 words were classified into two main categories, that is, the top 25 content words and function words; the content words were then further divided into nouns, adjectives, adverbs, and verbs as shown in Appendix 1, though some words may belong to more than one category. From Appendix 1, it can be seen that the verbs *have, take, has, make* and *stay* are the top five keyness-words, but would presumably not be keywords in other genres. Furthermore, a number of the adjectives used in the hotel brochures are past participles such as *designed, equipped, conditioned* or *located*, and sentences beginning with participle phrases are also usual in this genre. This usage implies that the passive voice is dominant throughout authentic hotel brochures as it often is in other domains of academic discourse and language registers. Upon closer examination it is found that these keywords may collocate with certain words mainly in this genre. Take the verb *enjoy* for instance; the usage of gerunds following it is surprisingly rare in these texts; only one such example is found in the 158 entries, that is, *Guests may enjoy dining on the sun terrace of….* Instead, nouns or noun phrases mostly follow the verb *enjoy* such as *enjoy all these great facilities, enjoy a drink* or *enjoy the unique flavour of*. An excerpt of its collocations extracted by the function concordance of WS is shown below in Table 4.

| N | Concordance |
|---|---|
| 1 | …of our clubs, you can join and *enjoy* all these great facilities… |
| 2 | …eating some me time, you'll *enjoy* free use of our Otium H.. |
| 3 | . …Or why not dine in style and *enjoy* superb food in the rela… |
| 4 | …refurbished, where guests can *enjoy* a drink or a bite to ea… |
| 5 | …advice you need to arrange and *enjoy* your special day. No st… |
| 6 | …ed to make it easier for you to *enjoy* the finer things in life… |
| 7 | …styled surroundings. Guests can *enjoy* fine dining or more rel… |
| 8 | …ave inter-connecting rooms and *enjoy* views over Hyde Park or… |
| 9 | …met. * Butler's Restaurant V *enjoy* the finest of English c… |
| 10 | …traditional afternoon tea and *enjoy* surroundings reminiscent of a… |
| 11 | …essentially English institution *Enjoy* a bewildering array of… |
| 12 | …Share a chat over coffee or *enjoy* an intimate dinner for… |
| 13 | …There's no better place to *enjoy* an English afternoon tea… |
| 14 | …Club InterContinental members *enjoy* a complimentary breakfast… |
| 15 | …NENTAL Guests of all 60 suites *enjoy* complimentary membership… |

**Table 4.** The partial concordance of the verb *enjoy* in authentic texts

### 4.3 The wordlists for students' texts

Tables 2 and 3 indicate that on average the students used 2,140 distinct words across the three versions of their texts, while 4,316 distinct words were identified in the AT. This is a difference of 2,176 words, and the number of total tokens in AT is also much larger than that in ST. Surprisingly, it is found that in fact the number of DW in every 100 tokens in AT is relatively fewer than that found in ST, that is, 14.46: 17.86[5]. This indicates that the students used more different or distinct words than those adopted in AT. ESP practitioners therefore need to highlight the words which are required in AT and address which words appearing in ST should be avoided.

**4.4 The 3 versions of keyword lists in the students' texts compared to the authentic texts**
As shown in Table 2, with reference to BNC, there are respectively 202, 266, and 270 words identified as keywords in ST, and the majority of these words overlap with the keywords in AT. This suggests that on the one hand these overlapping words are exactly the keywords that are used in hotel texts, but on the other hand the students used only about one third of the AT keywords. However, it would be more significant to examine which words are overused and underused by the students in terms of creating hotel brochure texts. Hence, the ST was used as the study corpus and the AT as the reference corpus to locate these words using WS.

In the end, three KW lists were generated as shown in Appendix 2. The shaded italic words are underused words, which should have been used more frequently, while the rest are overused words, which should not have appeared so often. The following are possible reasons to explain why some words are overused while others are underused.

The first reason is related to contextual differences, that is, the cultural factors existing in Taiwan and the UK, as well as micro personal experience. Following are some obvious examples. The students preferred using the word *buffet* in their brochures when introducing food and beverage (F&B) service. However, in the UK hotels *buffets* (all you can eat) are rare. Only 7 entries in AT are identified, which is only one-third of the number in ST. In fact, many Taiwanese go to hotels specifically to enjoy their buffets, and Taiwanese hotels also intentionally promote their F&B service by claiming the wide selection of dishes to attract guests. All-you-can-eat buffets are, indeed, indispensable for Taiwanese people when choosing a hotel, and this could be a strong cultural preference. Secondly, sometimes the students overused certain words simply because they had just experienced related activities and would like to include personal experiences in the texts to make them unique. Instances of this category include *chocolate*, *castle*, or *cat*. During the teacher-student feedback activities, it was found that some students had just come back from visiting a chocolate factory and a castle in the UK, and therefore they added these venues as the hotel excursion service. Besides, it could also be very likely that these students had experienced lodging in a hotel which offered chocolate on guests' arrival in their bedrooms. Hence, contextual factors and community constraints greatly affect textual conventions, and writers need to consider these boundaries (Johns, 2002). This situation may be improved if ESP learners are explicitly instructed on or become aware of the hotel cultures in a particular context.

The second reason is the fossilisation of English learning, meaning that Taiwanese ESP learners, as with ELT learners, tend to use formal English, especially a number of so-called elegant sentence patterns, to demonstrate their English proficiency. Therefore, the tendency to use the pattern *There is(are)…* means the words *there* and *are* become overused. Although it is grammatically acceptable to use these sentence patterns, they may make the language of the brochures either too formal or inauthentic. In other words, a pattern like this cannot convey the spirit of hospitality and friendliness. A sense of belonging is important for visitors in hotels, and appropriate word choices enhance this feeling. For example, the words *we* and *our*

[5] The number 17.86 is based on the average of 18.41, 17.85 and 17.32 of the students' three versions in DW per 100 tokens (see Table 3.).

are found more frequently than *you* and *your* in authentic hotel brochures (Yang, 2010a, 2010b).

Finally, the last reason may be the result of the users' unilateral lexical knowledge, which can explain most of the overused words in the three keyword lists. This indicates that learners have limited vocabulary and thus they tend to use one word to express a particular meaning rather than several. Usually, words have several meanings and uses, but students may learn only one. Therefore, they tend to continuously use certain words, which then become overused. These words can be further classified as vague and improper choices. The students tended to use certain verbs such as *provide* to refer to all the services a hotel has for guests. They also used the words *different* and *equipment* to include all the things they were unable to specify more clearly, such as *different styles, different ways, different scales, different languages, sound equipment, necessary equipment,* and *latest equipment*, which are seldom seen in AT. Rather than saying *different*, it is suggested that all the possibilities be listed to give guests a clear idea of their choices. For example, the sentence "*Food and drinks are provided with many different exotic cuisines in the hotel*" should be rewritten as "*The hotel offers Thai, Japanese, and Chinese cuisines as part of its F&B service.*" In addition, they also overused the adjectives *supreme* and *glamorous*. The second sub-category refers to the inappropriate choice of words. Words such as *visitors* and *customers* should be replaced by *guests*. The words, *meeting* and *conference* should be differentiated in terms of the scale of room size, catering service or facilities attached. In addition, the preposition *during* was overused in ST, but was not found at all in AT. The last example is the word *room*, which was overused to refer to the hotel guest rooms (e.g., *single room, standard room or triple room*) mostly in ST. However, in AT this word has more forms such as *living room, guest room, fitness room, meeting room, health room, event room,* and *tasting room*. Students use a limited number of terms and thus simply chose a vague or improper word while writing.

It is believed that this situation can also be gradually improved through explicit teaching of vocabulary. For underused words, teachers can address the various uses of each word and encourage learners to try using it in a more creative way, and with regards to the overused words students can be instructed in how to replace them with more specific or appropriate words which may appear either in the AT wordlist or keyword list.

**4.5 Evaluation of the three texts**

As Table 3 suggests, in ST the number of keywords (excluding special terminology) across the three stages gradually decreased, that is, from 20, to 16, and then 12, and the keywords in every 100 DWs was 1.07, 0.72 and 0.08. Apparently, the students' overused words gradually diminished and this decrease suggests that they tried to apply more of the AT keywords after the instructor's explicit intervention. Therefore, the project succeeded in reaching one of its initial aims, namely to help the learners write more authentic-flavoured texts. However, it should also be noted that compared to their V1 writing, the students wrote fewer words in their V2 and V3 texts. This may also be one reason to account for the fewer appearances of the keywords in ST.

The quantitative analysis above is not the only measurement of the students' progress. Qualitative measurement was also used to evaluate the students' writing. The following table displays the result of the English readers' assessment. Over half of the evaluators were native English speakers, and their perceptions of the three texts and nationality distribution are displayed in Tables 5 and 6.

| **Students' texts** | **The 1st Version** | | **The 2nd Version** | | **The 3rd Version** | |
|---|---|---|---|---|---|---|
| | **ANR**[6] | **MVV** | **ANR** | **MVV** | **ANR** | **MVV** |
| **Better version** | 4.16 | 4.16 | 4.16 | 4.16 | 91.66 | 91.66 |

**Table 5.** The percentages (%) of the English readers' choices of the better brochure

| | **Nationalities** | **Percentages** |
|---|---|---|
| English as the 1st language (***EL1***) | UK, USA, South Africa, and Australia. | 66.66% |
| English as a 2nd/foreign language (***EL2***) | Hungary, Thailand, Malaysia[7], Singapore, South Korea, Philippines, India, and Japan. | 33.33% |

**Table 6.** The nationality distribution and percentages of the English-speaking readers

Nearly all of the reviewers perceived that the third version was the best in terms of authenticity and motivation to pay a visit. The reason some early versions, V1 and V2, were used is because the evaluators regarded them as being as good as the third version, V3; therefore one version was sufficient. In most cases, it can be claimed that the last version was distinguished from the previous ones, confirming that the instruction was effective. After all, returning to the purpose of publishing an informative and promotional text for hotels, how to display correct information and explain incentive programs to potential guests, and at the same time increase their motivation, is the most crucial point. In their written comments, the reviewers discussed their preferences after reading the brochures, and noted that the students' texts were exactly what is aimed for and expected from a hotel brochure.

In sum, similar to Henry and Roseberry's (1998) research, this present study also confirms that explicit genre-based instruction does indeed improve students' writing. By analysing either the generic moves or linguistic features of a genre and then teaching them explicitly in class, ESP learners can become consciously aware of the constraints and choices allowed in a text which helps them achieve the specific communication purposes of the texts.

## 5 Implications and conclusion

Word frequencies and keyword lists are helpful in identifying essential words, grammar, and patterns in a special purpose area. In this study, 50 English hotel brochures together with the 24 students' self-designed texts were used as the samples to be analysed with Word Smith. In total, 602 keywords (KW) were identified in the authentic texts with reference to BNC; these words can play a significant role in English education for specific purposes in the hospitality

[6] ANR denotes authenticity; MVV denotes motivation to visit.
[7] English is spoken as one of the official languages in Singapore, Malaysia, Philippines and India.

and tourism industries, as they are the essential words required to compose effective promotional texts. In addition, this study analysed ESP learners' work in a GBI writing course. It was found that these ESP learners in Taiwan used different vocabulary when writing a specific genre text when compared with that used in authentic texts. The likely underlying reasons for this phenomenon may stem from three main influences: contextual differences/personal experience, fossilisation of English learning, and unilateral lexical knowledge. The decrease in overused keywords in the students' texts and the outsiders' positive evaluations both confirm that the students' third version was the best text in terms of its feeling of authenticity and the reader's desire to visit the hotel, which largely achieves the purposes of preparing an informational and promotional hotel brochure.

Words are one of the core elements in constructing a piece of writing. Choosing the right words and knowing their constraints help to construct the authenticity of a specific genre. The results of this research suggest that in each genre there are essential keywords which are frequently used, and ESP learners must learn to employ these words to write texts in that specific genre. Acquiring these words reduces inaccuracy and in-authenticity. Thus, it is hoped that ESP practitioners and learners can pay close attention to word frequencies and keyword lists for specific genres, pinpoint the authenticity and appropriateness of a particular genre, and disseminate these words in their professional community. As Paltridge (2001) argues, genre analysis provides ESP learners with not only knowledge and skills but also with access to socially powerful forms of language. Furthermore, presenting these lists in a form learners will read and use when writing will empower ESP educators to design curricula, materials, and activities, thus giving them a more central role (Hyland, 2004a, 2008).

Secondly, as Dudley-Evans (1995) reminds us, genre analysis or GBI should not become prescriptive but should be descriptive. It would be wrong for teachers to assert that many linguistic problems would be prevented or cured by analysing a genre and conducting GBI. What genre analysis can do is describe language use, not predict problems or cure them. Indeed, what a GBI classroom aims for is to equip ESP learners with abilities such as the generic competence (Bhatia, 2000) to "participate in and respond to new and recurring genres" (Paltridge, 2001: 7). Therefore, compared to EGP teaching, it may take teachers more time and patience to design and teach an ESP course, but depending on the methods of teaching EGP to teach ESP can be truly prescriptive and less effective.

Furthermore, it is suggested that ESP teachers teach the related grammar and contexts of the keywords. Rather than teaching general syntactic patterns, it is important to teach the grammar and context of the lexis and the lexical phrases. As shown in the data, some words may have totally different usage and meanings, and certain words are preferred in a particular genre. Genre-based instruction addresses this fact and raises learners' awareness of the allowable choices and constraints of a genre text.

In other words, researching and instructing the keywords of a particular genre can help teachers and learners perceive the *keyness*, *aboutness* and cultures of the words used in that genre. Using them appropriately and in a timely manner can allow prospective stakeholders to communicate with each other without confronting inauthentic, ambiguous or misunderstood situations in a professional community.

To conclude, this research is a preliminary investigation of an area that deserves further attention in Taiwan, and which can serve as a starting point for further research. In future studies, researchers could investigate how the keywords are used in website texts, and brochure and web text could be compared. In addition, the difference in keyword usage between experimental and control groups can be compared to investigate the impact of keyword instruction. Since all of the authentic texts used in this study came from the UK, the research would be more representative if authentic texts from other English contexts could be included. Besides, in the E-era an increasing number of hotels have stopped publishing printed brochures and are putting all their information on their websites instead, as the Internet can offer much more space, information and cheaper prices in comparison with printed texts. Will it be likely that the choices and constraints on words are different owing to the expansion of word storage space? And will this lead to wordlists and keywords which differ from those in this paper when compared with BNC or other corpora? This issue is indeed worthy of further investigation.

## 6 References

Archer, D. (2009) (ed.): *What's in a Word-list? Investigating word frequency and keyword extraction.* Surrey, England, UK: Ashgate.

Archer, D. Culpeper, J. & Rayson, P. (2009): Love-'a familiar or a devil'? An exploration of Key Domains in Shakespeare's Comedies and Tragedies. In D. Archer (ed.): *What's in a Word-list? Investigating word frequency and keyword extraction* (pp. 137-157). Surrey, England, UK: Ashgate.

Baker, P. (2009): 'The question is, how cruel is it?' Keywords, Fox Hunting and the House of Commons. In D. Archer (ed.): *What's in a Word-list? Investigating word frequency and keyword extraction* (pp. 125-136). Surrey, England, UK: Ashgate.

Biber, D. & Conrad, S. (2009): *Register, genre, and style.* New York, NY: Cambridge University Press.

Bhatia, V. K. (1993): *Analysing genre: Language use in professional settings.* London: Longman.

Bhatia, V. K. (1997): Applied Genre Analysis and ESP. In T. Miller (ed.): *Functional Approaches to Written Text: Classroom Applications* (pp.134-149). Washington D.C.: US Information Agency.

Bhatia, V. K. (2000): Integrating discursive competence and professional practice: A new challenge for ESP. Paper presented at the *TESOL colloquium "Rethinking ESP for the New Century,"* Vancouver, March.

Carstens, A. (2008). A multidirectional model for tertiary-level disciplinary writing. *Journal for Language Teaching,* 42(1): 81-97.

Cheng, A. (2006): Understanding learners and learning in ESP genre-based writing instruction. *English for Specific Purposes,* 25(1): 76-89.

Crystal, D. (1991): *A dictionary of linguistics and phonetics*. Oxford, UL: Basil Blackwell.

Davis, M. (2009): Word frequency in context: alternative architectures for examining related words, register variation and historical change. In D. Archer (ed.): *What's in a Word-list? Investigating word frequency and keyword extraction* (pp. 53-68). Surrey, England, UK: Ashgate.

Dudley-Evans, T. (1995): Genre models for the teaching of academic writing to second language speakers: Advantages and disadvantages. *Journal of TESOL France,* 2(2): 181-192.

Ellis, R. (1995): Interpretation tasks for grammar teaching. *TESOL Quarterly,* 29(1): 87-105.

Hafner, C. (2010): A Multi-perspective Genre analysis of the Barrister's Opinion: Writing Context. *Generic Structure and Textualisation,* 27(4): 410-441.

Halliday, M. A. K., Mclntosh, A. & Strevens, P. (1964). *The linguistics sciences and language teaching*. London: Longman.
Henry, A. & Roseberry, R. (1996): A Corpus-based Investigation of the Language and Linguistics Patterns of One Genre and the Implications for Language Teaching. *Research in the Teaching of English*, 30(4): 472-489.
Henry, A. & Roseberry, R. (1998): An Evaluation of a Genre-Based Approach to the Teaching of EAP/ESP Writing. *TESOL Quarterly*, 32(1): 147-156.
Henry, A. & Roseberry, R. (2007): Language Errors in the Genre-based Writing of Advanced Academic ESL Students. *RELC Journal*, 38(2): 171-198.
Hyland, K. (2004a): *Genre and Second Language Writing*. Ann Arbour, MI: University of Michigan Press.
Hyland, K. (2004b): Graduates' gratitude: the generic structure of dissertation acknowledgements. *English for Specific Purposes*, 23(3): 303-324.
Hyland, K. (2007): Genre pedagogy: language, literacy and L2 writing instruction. *Journal of Second Language Writing*, 16(3): 148-164.
Hyland, K. (2008): As can be seen: Lexical bundles and disciplinary variation. *English for Specific Purposes*, 27(1): 4-21.
Hyland, K. (2010): Community and Individuality: performing Identity in Applied Linguistics, *Written Communication*, 27(2): 159-188.
Hyland, K. & Tse, P. (2007): Is there an 'academic vocabulary'? *TESOL Quarterly*, 41(2): 235-253.
Hyon, S. (2008): Convention and inventiveness in an occluded academic genre: A case study of retention-promotion-tenure reports. *English for Specific Purposes*, 27(2): 175-192.
Johns, A. (1995): Teaching classroom and authentic genres: Initiating students into academic cultures and discourses. In D. Belcher & G. Braine (eds.): *Academic writing in a second language: Essays on research and pedagogy* (pp. 277-291). Norwood, NJ: Ablex.
Johns, A. (1997): *Text, role, and context: Developing academic literacies*. Cambridge: Cambridge University Press.
Johns, A. (ed.) (2002): *Genre in the classroom: Multiple perspectives*. Mahwah, NJ: Erlbaum.
Johns, A. (2007): *Genre awareness for the novice academic student: An on-going quest*. Paper presented in the American Association of Applied Linguistics Conference, Costa Mesa, California.
Karlsson, A. (2009): Positioned by Reading and Writing: Literacy Practices, Roles, and Genres in Common Occupations. *Written Communication*, 26(1): 53-76.
Kirk, J. M. (2009): Word Frequency: use or misuse? In D. Archer (ed.): *What's in a Word-list? Investigating word frequency and keyword extraction* (pp.17-34). Surrey, England, UK: Ashgate.
Lam, P. (2006): "Tree-lined or tree-fringed?" A corpus-driven approach to the study of the adjectival compounds in English travel and tourism industry texts. *The 3rd IVACS group International Conference on Language at the Interface*. University of Nottingham, Nottingham, U.K., 23-23 June 2006.
Lee, D. Y. W. (2001): Genres, registers, text types, domains, and styles: Clarifying the concepts and navigating a path through the BNC jungle. *Language Learning & Technology*, *5*(3): 37-72.
Lee, D. Y. W. & Chen, S. X. (2009): Making a bigger deal of the smaller words: Function words and other key items in research writing by Chinese learners. *Journal of Second Language Writing*, 18(3): 281-296.
Lien, Y. L. (2008): *A Coupus-based genre analysis of English promotional travel texts*. Unpublished master dissertation. Taoyuan: Ming-chuan University, Taiwan.

Maingueneau, D. (2002): Analysis of an academic genre. *Discourse Studies*, 4(3): 319-342.

Martin, J. R. (1993): A contextual theory of language. In B. Cope & M. Kalantzis (eds.): *The powers of literacy: A genre approach to teaching writing* (pp. 116-136). London: Falmer Press.

Matt, H. (2007): How promotional language in press releases is dealt with by journalists: Genre mixing or genre conflict? *Journal of Business Communication*, 44(1): 59-95.

Moya Guijarro, A. J. (2006): The continuity of topics in journal and travel texts: A discourse of functional perspective. *Functions of Language*, 13(1): 37-76.

Moya Guijarro, A. J. & Albentosa Hernandez, J. I. (2001): Points of departure in news items and tourist brochures: Choices of theme and topic. *Text, 21*(3): 347-371.

Osman, H. (2008): Re-branding academic institutions with corporate advertising: genre perspective. *Discourse & Communication,* 2(1): 57-77.

Paltridge, B. (2001): *Genre and the Language Learning Classroom.* Ann Arbour, MI: University of Michigan Press.

Samraj. B. & Monk, L. (2008): The statement of purpose in graduate program applications: Genre structure and disciplinary variation. *Journal of English for Specific Purposes,* 27(2): 193-211.

Schneider, B. & Andre, J. (2005): University preparation for workplace writing: An exploratory study of the perceptions of students in three disciplines. *Journal of Business Communication*, 42(2): 195-218.

Scott, M. (2008): WordSmith Tools (Version 5.0) [Computer Software]. Liverpool, UK: Lexical Analysis Software.

Scott, M. & Tribble, C. (2006): *Textual Patterns: Keywords and Corpus Analysis in Language Education.* Amsterdam: Benjamins.

Swales, J. M. (1990): *Genre analysis: English in academic and research settings.* Cambridge: Cambridge University Press.

Ventola, E. (1995): Generic and register qualities of texts and their realisation. In H. Peter & M. Gregory (eds.): *Discourse in society: Systemic functional perspectives* (pp. 3-28). Norwood, NJ: Ablex.

Vergaro, C. (2004): Discourse strategies of Italian and English sales promotion letters. *English for Specific Purposes,* 23(2): 181-207.

Wennerstrom, A. (2003): *Genre analysis in the language classroom. Vol.2, Genres of Writing.* Ann Arbour, MI: University of Michigan Press.

Wildes, V. & Nyheim, P. (2009): *Business Writing for Hospitality.* Upper Saddle River, NJ: Person Education.

Woodward-Kron, R. (2005): The role of genre and embedded genres ion tertiary students' writing. *Prospect,* 20(3): 24-41.

Wrigglesworth, J. & M. McKeever. (2010): Writing history: A genre-based, interdisciplinary approach linking disciplines, language, and academic skills. *Arts and Humanities in Higher Education*, 9(1), 107-126.

Yang, W. H. (2010a): Analysing Keywords in English Hotel Brochures. *The Proceedings of 2010 International Symposium on ESP and Its Implications in Nursing & Medical English Education* (pp. 298-313). Kaohsiung: Fooyin University, Taiwan.

Yang, W. H. (2010b): One genre in briefs: A Computational analysis of word-listing hospitality English in hotel information brochures. *The Proceedings of 2010 International Conference on ELT Technological Industry and Book Fair* (pp. 120-135). Pingtung: National Pingtung University of Science & Technology, Taiwan.

Zhu, Y. (2005): *Written communication across cultures: A sociocognitive perspective on business genres.* Amsterdam: Benjamins.

## 10 Appendix

**Appendix 1: Top 25 Categorised Content and Functional Words in the UK Text**

| | **Verbs** | **Nouns** | **Adjectives** | **Adverbs** | **Functional Words** |
|---|---|---|---|---|---|
| 1 | Have | # | Our | more | the |
| 2 | Take | hotel | Your | also | and |
| 3 | Has | you | All | most | a |
| 4 | Make | rooms | One | out | to |
| 5 | Stay | room | High | just | of |
| 6 | Enjoy | we | Every | there | in |
| 7 | Walk | bar | Available | well | with |
| 8 | Offer | that | New | where | for |
| 9 | Offers | hotels | Great | then | be* (is) |
| 10 | Visit | it | Modern | some | at |
| 11 | Work | meeting | Perfect | away | from |
| 12 | Find | business | Free | fully | or |
| 13 | Design | service | Contemporary | not | on |
| 14 | Choose | guests | International | even | be*(are) |
| 15 | Need | suites | Stylish | simply | an |
| 16 | Welcome | bedrooms | Located | only | by |
| 17 | Relax | suite | Which | along | can |
| 18 | End | restaurant | No | truly | as |
| 19 | Use | facilities | Special | again | be |
| 20 | Call | dining | Complimentary | indoor | will |
| 21 | Exit | park | First | very | up |
| 22 | Escape | city | Royal | across | onto |
| 23 | Feel | day | Unique | next | over |
| 24 | See | centre | Designed | here | within |
| 25 | Turn | street | Two | nearby | towards |

(*from Yang, 2010a*)

## Appendix 2: Three keyword lists in ST with reference to AT

**The keywords in students' 1st self-designed texts V1 compared to the authentic texts**

| N | Keyword | Freq. | % | RC. Freq. | RC. % | Keyness | P |
|---|---|---|---|---|---|---|---|
| 1 | BIRMINGHAM | 98 | 0.945 | 47 | 0.1521 | 115.75 | 1E-15 |
| 2 | CONFERENCE | 70 | 0.675 | 52 | 0.1683 | 57.219 | 6E-14 |
| 3 | CUSTOMERS | 22 | 0.2122 | 1 | | 53.166 | 1E-13 |
| 4 | EQUIPMENT | 38 | 0.3664 | 14 | 0.0453 | 52.586 | 1E-13 |
| 5 | ARE | 112 | 1.08 | 145 | 0.4692 | 41.595 | 2E-12 |
| 6 | THERE | 45 | 0.4339 | 31 | 0.1003 | 39.583 | 5E-12 |
| 7 | PROJECTOR | 14 | 0.135 | 0 | | 38.692 | 8E-12 |
| 8 | VISITORS | 17 | 0.1639 | 2 | | 35.356 | 3E-10 |
| 9 | PROVIDE | 27 | 0.2604 | 11 | 0.0356 | 35.271 | 4E-10 |
| 10 | TWO | 39 | 0.3761 | 28 | 0.0906 | 32.945 | 7E-09 |
| 11 | DIFFERENT | 23 | 0.2218 | 8 | 0.0259 | 32.798 | 7E-09 |
| 12 | CAN | 78 | 0.7522 | 99 | 0.3203 | 30.049 | 4E-08 |
| 13 | SUPREME | 14 | 0.135 | 2 | | 27.792 | 1E-07 |
| 14 | TUDOR | 14 | 0.135 | 2 | | 27.792 | 1E-07 |
| 15 | CHOCOLATE | 12 | 0.1157 | 1 | | 26.69 | 2E-07 |
| 16 | DURING | 17 | 0.1639 | 5 | 0.0162 | 26.294 | 3E-07 |
| 17 | CASTLE | 14 | 0.135 | 3 | | 24.582 | 7E-07 |
| 18 | GLAMOROUS | 11 | 0.1061 | 1 | | 24.093 | 9E-07 |
| *19* | *#* | *261* | *2.5169* | *1076* | *3.4815* | *-24.32* | *8E-07* |
| *20* | *HOTELS* | *5* | *0.0482* | *98* | *0.3171* | *-30.57* | *3E-08* |

**The keywords in students' 2nd self-designed texts V2 compared to the authentic texts**

| N | Keyword | Freq. | % | RC. Freq. | RC. % | Keyness | P |
|---|---|---|---|---|---|---|---|
| 1 | BIRMINGHAM | 75 | 0.580900013 | 47 | 0.152074024 | 53.62804413 | 1.05793E-13 |
| 2 | CAN | 105 | 0.813260019 | 99 | 0.320326149 | 43.31893539 | 9.88533E-13 |
| 3 | EQUIPMENT | 36 | 0.278832018 | 14 | 0.045298647 | 38.50824738 | 8.73106E-12 |
| 4 | PROVIDE | 31 | 0.240105331 | 11 | 0.035591796 | 35.1747818 | 4.93422E-10 |
| 5 | CUSTOMERS | 17 | 0.131670669 | 1 | | 34.53507996 | 1.26168E-09 |
| 6 | PROJECTOR | 14 | 0.10843467 | 0 | | 34.22509384 | 1.98448E-09 |
| 7 | TAIWAN | 14 | 0.10843467 | 0 | | 34.22509384 | 1.98448E-09 |
| 8 | KAOHSIUNG | 13 | 0.100689337 | 0 | | 31.77973366 | 1.43438E-08 |
| 9 | ARE | 122 | 0.944930673 | 145 | 0.46916455 | 31.43531227 | 1.7695E-08 |
| 10 | DIFFERENT | 25 | 0.193633333 | 8 | 0.025884941 | 30.15326309 | 3.69968E-08 |
| 11 | THERE | 45 | 0.348540008 | 31 | 0.100304149 | 28.90882874 | 7.29416E-08 |
| 12 | BUFFET | 23 | 0.178142667 | 7 | 0.022649324 | 28.52201462 | 8.97145E-08 |
| 13 | CONFERENCE | 60 | 0.464720011 | 52 | 0.168252125 | 28.32373047 | 9.97054E-08 |
| 14 | ROOM | 113 | 0.875222683 | 136 | 0.440044016 | 28.22018051 | 1.05345E-07 |
| 15 | ƒÜ | 11 | 0.085198671 | 0 | | 26.88934135 | 2.12519E-07 |
| 16 | DELETE | 11 | 0.085198671 | 0 | | 26.88934135 | 2.12519E-07 |
| 17 | TOUR | 15 | 0.116180003 | 2 | | 25.75036049 | 3.8563E-07 |
| 18 | DURING | 19 | 0.147161335 | 5 | 0.016178088 | 25.37663078 | 4.68672E-07 |
| *19* | *HOTELS* | *5* | *0.038726669* | *98* | *0.317090541* | *-40.70198059* | *2.65754E-12* |
| *20* | *O* | *3* | *0.023236001* | *87* | *0.28149873* | *-41.82489014* | *1.66976E-12* |
| *21* | *S* | *21* | *0.162652001* | *223* | *0.721542716* | *-64.14376068* | *2.92077E-14* |

**The keywords in students' 3rd self-designed texts V3 compared to the authentic texts**

| N | Keyword | Freq. | % | RC. Freq. | RC. % | Keyness | P |
|---|---|---|---|---|---|---|---|
| 1 | BIRMINGHAM | 81 | 0.5793 | 47 | 0.1521 | 55.905 | 8E-14 |
| 2 | TAIWAN | 18 | 0.1287 | 0 | | 42.005 | 2E-12 |
| 3 | KAOHSIUNG | 15 | 0.1073 | 0 | | 35.002 | 8E-10 |
| 4 | WE | 115 | 0.8224 | 119 | 0.385 | 32.953 | 7E-09 |
| 5 | CAN | 99 | 0.708 | 99 | 0.3203 | 30.499 | 3E-08 |
| 6 | SUPREME | 18 | 0.1287 | 2 | | 30.493 | 3E-08 |
| 8 | PROVIDE | 29 | 0.2074 | 11 | 0.0356 | 28.835 | 8E-08 |
| 9 | BUFFET | 24 | 0.1716 | 7 | 0.0226 | 28.114 | 1E-07 |
| 10 | PROJECTOR | 12 | 0.0858 | 0 | | 28 | 1E-07 |
| 11 | CAT | 12 | 0.0858 | 0 | | 28 | 1E-07 |
| 12 | GARFIELD | 12 | 0.0858 | 0 | | 28 | 1E-07 |
| 13 | ARE | 126 | 0.9011 | 145 | 0.4692 | 27.989 | 1E-07 |
| 14 | DURING | 21 | 0.1502 | 5 | 0.0162 | 27.28 | 2E-07 |
| 15 | ROOM | 118 | 0.8439 | 136 | 0.44 | 26.097 | 3E-07 |
| *17* | *HOTELS* | *5* | *0.0358* | *98* | *0.3171* | *-44.89* | *6E-13* |
| *18* | *HOTEL'S* | *3* | *0.0215* | *98* | *0.3171* | *-53.23* | *1E-13* |

***

LSP Professional Communication Knowledge Management Cognition

LSP Journal, Vol.3, No.1 (2012) / http://lsp.cbs.dk

# The configuration of a philosophical parameter in the subontology #ENTITY of FunGramKB: The case of axiology[1]

Ángel Felices-Lago [a], María Enriqueta Cortés-de-los-Ríos [b], Amalia Marín-Rubiales [c]

[a] Department of English & German
University of Granada, Spain
afelices@ugr.es

[b] Department of English & German
University of Almería, Spain
mecortes@ual.es

[c] Department of English & German
University of Córdoba, Spain
ff1marua@uco.es

***Keywords:*** *ontology, terminology, axiology, FunGramKB, Knowledge base*

**Abstract**

The Functional Grammar Knowledge Base *FunGramKB* (FGKB), on the one hand, is a multipurpose lexico-conceptual knowledge base for natural language processing (NLP) systems. It comprises three major interrelated knowledge level modules: lexical, grammatical and conceptual. At the conceptual level the core ontology is presented as a hierarchical catalogue of the concepts that a person has in mind. Here is where semantic knowledge is stored in the form of meaning postulates. On the other hand, axiology is considered to be a primitive, basic or key parameter, among others, in the architecture of meaning construction at different levels. This parameter can be traced back to the three subontologies in which FunGramKB can be split: #ENTITY for nouns, # EVENT for verbs, and #QUALITY for adjectives. In this paper we shall concentrate on the category # ENTITY and explore how the main categories and features of the axiological parameter (good-bad or positive-negative [+/-]) are represented and encoded within FunGramKB ontology, particularly inside semantic properties such as the meaning postulates.

## 1 Basic assumptions

In this study we start from two premises: the first one states that *valuation* is an inherent aspect of categorization. In fact, in the ontogenetic development of every human being, the first categorizations are valuations. The reason is that we are valuating beings. All our actions, our thinking, our attitudes and interactions with the world and with other people, but

[1] Financial support for this research has been provided by the DGI, Spanish Ministry of Economy and Competitiveness, grant FFI2010-15983.

particularly our emotions are connected with or laden with certain values (Krzeszowski, 1997). Consequently, axiology is considered to be a primitive, basic or key parameter, among others, in the architecture of meaning construction at different levels in language (Hare 1952; Osgood, Suci, Tannenbaum 1957; Coseriu 1967; Pottier 1974; Krzeszowski 1990, 1993, 1997; Felices-Lago 1991, 1997; Cortés de los Ríos 2001; and many others).

In the second place, in the last few years the comprehensive theory of constructional meaning known as the Lexical Constructional Model (Mairal-Usón & Ruiz-de-Mendoza-Ibáñez 2008, 2009; Ruiz-de-Mendoza-Ibáñez & Mairal-Usón 2008, among others) has incorporated as part of its architecture *FunGramKB* (FGKB), which is a multipurpose lexico-conceptual knowledge base for natural language processing (NLP) systems (Periñán-Pascual & Arcas-Túnez 2004; Periñán-Pascual & Arcas-Túnez 2005; Mairal-Usón & Periñán-Pascual 2009; Mairal-Usón & Periñán-Pascual 2010; Periñán-Pascual & Mairal-Usón 2009; Periñán-Pascual & Mairal-Usón 2010). It is multipurpose in the sense that it is both multifunctional and multilingual. In other words, FGKB can be reused in various NLP tasks (e.g. information retrieval and extraction, machine translation, dialogue-based systems, etc.) and with several natural languages. This knowledge base comprises three major knowledge levels, consisting of several independent but interrelated modules: (1) Lexical level: *The Lexicon* stores morphosyntactic, pragmatic and collocational information about words. *The Morphicon* helps our system to handle cases of inflectional morphology. (2) Grammatical level: *The Grammaticon* stores the constructional schemata which take part in the bidirectional linking algorithm: semantics <-> syntax. (3) Conceptual level: *The Ontology* is presented as a hierarchical catalogue of the concepts describing semantic knowledge.[2] *The Cognicon* stores procedural knowledge by means of script-like schemata in which a sequence of stereotypical actions is organised on the basis of temporal continuity. *The Onomasticon* stores information about instances of entities and events. In FunGramKB, every lexical or grammatical module is language-dependent, whereas every conceptual module is shared by all languages. FunGramKB adopts a conceptualist approach to language, where the ontology becomes the pivotal module for the whole architecture.

As a consequence of the two previous premises, the valuation or axiological parameter can be traced back to the three subontologies in which the FunGramKB ONTOLOGY can be split: #ENTITY for nouns, #EVENT for verbs, and #QUALITY for adjectives (and some adverbs). In this paper we shall concentrate on the subontology #EVENT and explore how the main categories and features of the axiological parameter (good-bad or positive-negative [+/-]) are represented and encoded within the FunGramKB ontology. To do that, we should understand first how this ontology works on the basis of the following protocol: the FGKB Ontology stores semantic Knowledge in the form of thematic frames (TFs) and meaning postulates (MPs) by presenting a hierarchical catalogue of all the concepts (not ´words`, unlike FrameNet or MultiWordNet) that a person has in mind and works with two reasoning mechanisms: inheritance and inference, due to the fact that it is constructed on the basis of a deep semantic approach which not only displays concepts, but also defines them through a machine-readable metalanguage called COREL (i.e. *Conceptual Representation Language*).

Within each one of the three subontologies, FGKB also distinguishes three categories of concepts organized hierarchically:

[2] The FGKB Core ontology is deemed as an IS-A conceptual hierarchy which allows non-monotonic multiple inheritance. This ontology is both universal and linguistically-motivated.

*(a) Metaconcepts* (e.g. #PHYSICAL, #COLLECTION, #EMOTION, #TEMPORAL, etc.), which form the upper level in the taxonomy, as a result of the analysis of the most relevant linguistic ontologies, i.e. DOLCE, SIMPLE, SUMO, etc.

*(b) Basic concepts*, preceded by symbol +, are used as defining units which enable the construction of MPs for basic concepts and terminals, as well as taking part as selection preferences in TFs: e.g. +BOOK_00, +DIRTY_00, +FORGET_00, etc. They can be employed to define any word in any of the European languages that are claimed to be part of the Ontology. The starting point for the identification of basic concepts was the defining vocabulary in *Longman Dictionary of Contemporary English* (Procter 1978), though deep revision was required in order to perform the cognitive mapping into a single inventory of about 1,300 basic concepts.

*(c) Terminal concepts*, which are headed by symbol $. Terminals are not hierarchically structured and do not have definitory potential to take part in MPs: e.g. $METEORITE_00, $VARNISH_00, $CADAVEROUS_00, etc.

Basic and terminal concepts in FGKB are provided with semantic properties which are captured by *thematic frames* and *meaning postulates*. Every event in the ontology is assigned one single thematic frame, i.e. a conceptual construct which states the number and type of participants involved in the prototypical cognitive situation portrayed by the event (Periñán-Pascual & Arcas-Túnez 2007). Moreover, a meaning postulate is a set of one or more logically connected predications ($e_1$, $e_2$, … $e_n$), i.e. conceptual constructs that represent the generic features of concepts. As stated above, the basic concepts are the main building blocks of these types of constructs in the core ontology.

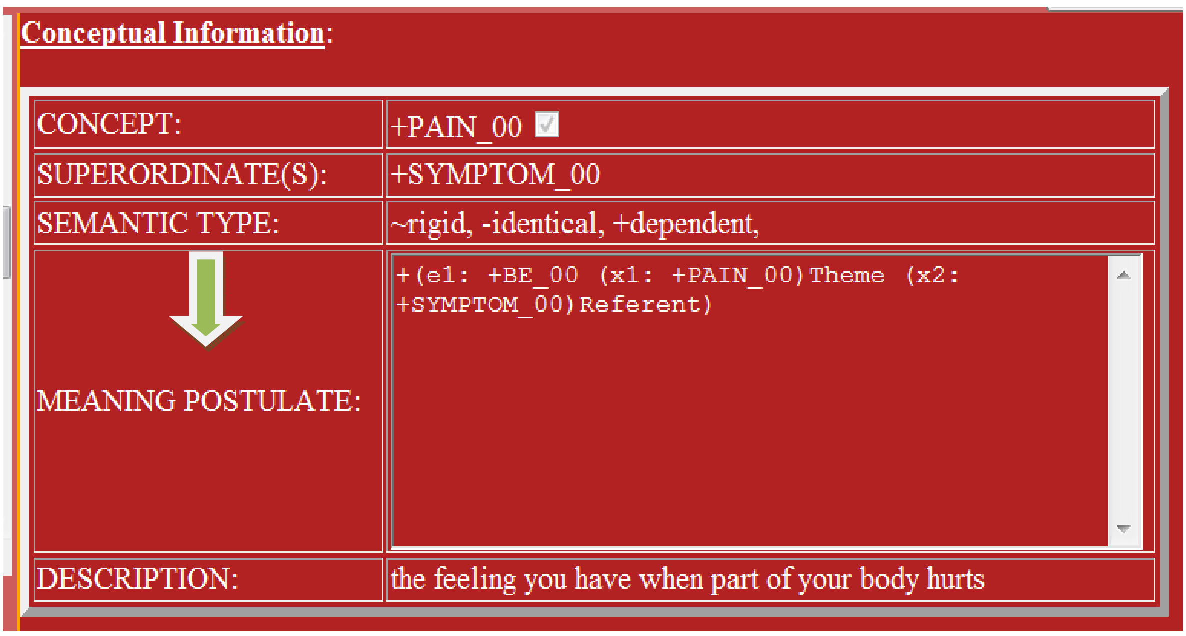



**Figure 1.** Meaning postulate of +PAIN_00 in FunGramKB

## 2 The axiological axis in the nominal lexicon

In the Functional Lexematic Model (FLM), Martín-Mingorance (1987), inspired by Coseriu (1967), introduced the category CLASSEMES, which were defined as general pragmatic, semantic and syntactic determinations in the vocabulary or as a kind of grammar. He proposed some of them for the pragmatic module:

- focus
- speaker`s evaluation
- social norm, etc.
- aesthetic norm.

Faber and Mairal-Usón (1999) proposed four macro-organizational patterns which appear across a wide range of verbal domains: Space; Time; Sociocultural context and Axiological evaluation (positive/negative). The first two patterns basically affect verbs, but the last two are shared by verbs, nouns or adjectives. The axiological pattern basically referred to Krzeszowski`s Lakoffian approach based on a the three-level hierarchy of values (sensory experience, life and health, spiritual level) given by classical axiologists such as Max Scheler or Józef Tischner. However, our proposal of axiological evaluation is based on a series of axes and scales that contribute to outlining the prototypical features characterizing its structure.

The first axis is preconceptual, lexico-genesic and dual, referring to its polar nature:

Positive (+) *vs.* Negative (-)

The second axis refers to the varying degrees of positiveness or negativity that are essential to domains or lexemes:

Maximum-Medium-Low ___ Low-Medium-Maximum

Positivity (+) Neutral (0) Negativity (-)

The third axis proposes a Hierarchy of axiological dimensions at intralinguistic level which are a result of an intensive research on lexicographical sources (Felices-Lago 1997: 188).

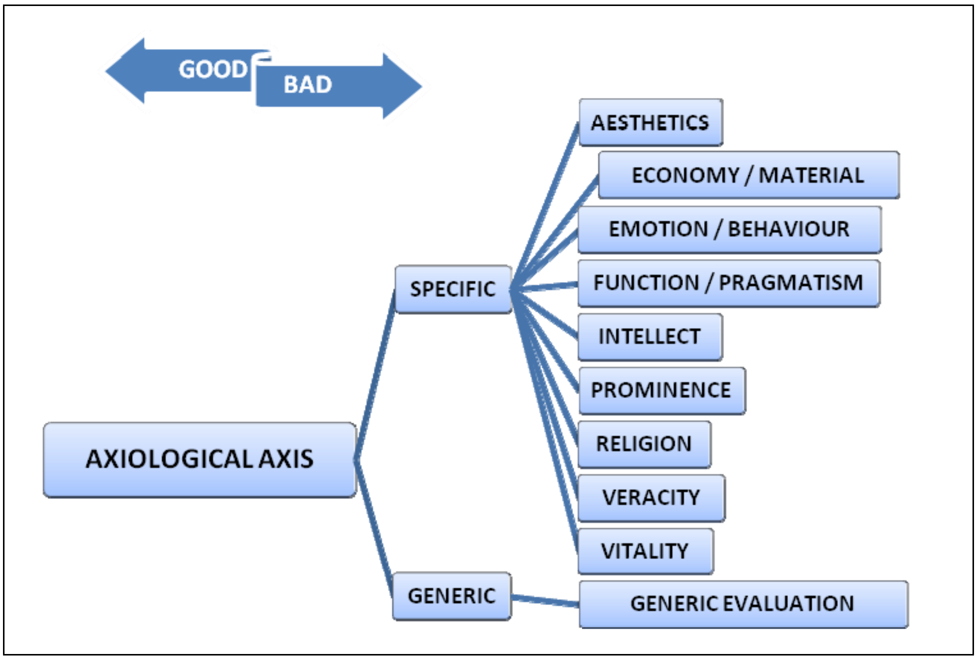



**Figure 2.** Hierarchy of axiological dimensions at linguistic level

## 3. Axiological representation and distribution in FGKB core ontology

Velardi et al. (1991) distinguish two well-defined strategies when describing meaning in computational lexicography: i.e. the cognitive content in a lexical unit can be described by means of semantic features or primitives (conceptual meaning), or through associations with other units in the lexicon (relational meaning). The former approach offers a stronger inferential power and guarantees the construction of a robust knowledge base applicable to most NLP tasks, consolidating thus the concept of resource reuse.

In FGKB, the meaning postulate (MP) is conceived as a property of basic concepts and terminals. Periñán-Pascual and Arcas-Túnez (2004) point out that current lexicalist models agree to handle lexical meaning as a cognitive representation reflecting the speakers` shared knowledge about the referent linked to a given linguistic expression. Therefore, when representing one of the meanings of a lexical unit, we are really representing the meaning of a concept. In consequence, an MP is a set of one or more logically connected predications, which are cognitive constructs carrying the generic features of the concept. If we apply a syntactico-semantic description to the participants, then a set of operators allows the machine to recognize well-formed predications.

If we link FGKB and the axiological parameter, in the following lines it will be observed how the axiological features are expanded and distributed throughout a set of semantic/ conceptual instruments (basic and terminal concepts, predications or satellites) and syntactic ones (predication operators such as polarity, quantification operators and logical connectors), in line with the process of stepwise conceptual decomposition characterizing FGKB.

### 3.1 Syntactic features of MPs: Operators

***Quantification Operators:***

A participant can be preceded by an operator (α), which applies a specific kind of quantification to the concept expressed as a selection preference.

| Feature | Value |
|---|---|
| Absolute quantifier | 1 / 2 / 3 / 4 … |
| **Relative quantifier** | **m / s / p** |
| Indefinite quantifier | i |

Ex: … (x7: **m** +GOOD_00)Attribute: **+PRIDE_00**

**Figure 3.** FunGramKB quantification operators

The quantification operators sensitive to axiological concepts are the relative quantifiers, particularly m, as it acts within the gradable semantic dimensions.

***Predication operators:***

Polarity operator *n* allows negative information to be explicitly stated. It is similar to *neg* in d-Prolog proposed by Nute (2003).

| **Feature** | **Value** |
|---|---|
| Aspectuality | ing / pro / egr |
| Temporality | rpast/npast /pres/nfut/rfut |
| Epistemic modality | cert / prob / pos |
| Non-epistemic modality | obl / adv / perm |
| **Polarity** | **n** |

Ex: +(e2: **n** +BE_01 (x1)Theme (x3: +LEGAL_00)Attribute): **+CRIME_00**

**Figure 4.** FunGramKB predication operators

### 3.2 Conceptual features of MPs: Predications and satellites

Only **basic concepts** can be used in Meaning Postulates to define terminal concepts or other basic concepts. A sample of axiologically-loaded basic concepts used in the meaning postulates of relevant units are shown as follows:

+IMPORTANT_00; +BEAUTY_00; +PLEASANT_00; +INTELLIGENCE_00; +WISDOM_00; +COURAGE_00; +CRUELTY_00; +DISLIKE_00; +DESIRE_00; +DESIRE_01; +FEAR_00; +PLEASURE_00; +LIKE_00; +PRIDE_00; +GOOD_00;+SADNESS_00; +SORROW_00; +DIE_00; +GOD_00; +WRONG_00; +FUNNY_00; +STUPID_00; +HAPPY_00; m +BAD_00; +COMFORT_00; +DANGER_00; +DAMAGE_00; +DAMAGE_01, etc.

These defining units that enable the construction of meaning postulates are limited to an inventory of about 1,300 units, which come mostly from defining vocabulary in *Longman Dictionary of Contemporary English*. They can be found in predications or satellites as it can be seen below:

***Predications:***

… (e5: +BE_01 (x5)Theme (x7: m +GOOD_00)Attribute) Referent: **+PRIDE_00**
*(e3: +BE_01 (x1)Theme (x3: +BEAUTIFUL_00)Attribute): **+FLOWER_00**
+(e2: n +BE_01 (x1)Theme (x3: +LEGAL_00)Attribute): **+ CRIME_00**
(e5: +DISLIKE_00 (x7)Agent (x4)Theme): **+ DISLIKE_00**

***Satellites:***

(… (f1: x1)Manner)(e3: +BE_01 (x4)Theme (x5: +BAD_00)Attribute): **+ EVIL_00**
(f2: (e4: +BE_01 (x3)Theme (x7: +GOOD_00)Attribute)): **+HEAVEN_00**
(f1: (e3: +BE_01 (x4)Theme (x5: +DIFFICULT_00)Attribute)): **+PROBLEM_00**
(f1: (e4: +BE_01 (x4)Theme (x8: +HAPPY_00 | +NERVOUS_00) Attrib)): **+DRUG_00**

### 4 Results and conclusions

For reasons of space we can only offer a sample of the entire analysis carried out in order to assess the impact of the axiological classeme in the FunGramKB core Ontology. The most relevant features can be summarized as follows:

LSP Journal, Vol.3, No.1 (2012) / http://lsp.cbs.dk

1. There is no reason to consider #ENTITIES as less sensitive to the axiological parameter than #EVENTS or #QUALITIES, even though only 74 out of 929 #ENTITIES are sensitive to inherent axiological information in their MPs (8%) and are distributed among the two leading metaconcepts like this:

   #ABSTRACT 38;
   #PHYSISCAL 36.

2. +DAMAGE_00 and +DAMAGE_01 are the most recurrent axiologically-loaded basic concepts: 6 times. They are followed by +PLEASANT_00, 4 times, +PLEASURE_00 and also +DISLIKE_00 3 times each. But considering that +PLEASANT_00 and +PLEASURE_00 refer to the same concept (as a quality or as an entity), it can be maintained that the hedonic concept is the most prominent here, as they both account for 7 occurrences.

3. The concepts from the core ontology in the third axis or hierarchy of axiologically-loaded dimensions can be distributed as follows:

   **A) GENERIC AXIS** -
   *PROTOTYPICAL EVALUATIVE CONCEPTS:*
   +GOOD_00, m +GOOD_00, m +BAD_00.

   **B) SPECIFIC AXIS**

   ***1a) EMOTION/BEHAVIOUR:***
   + PLEASANT_00, DESIRE_00, DESIRE_01, FEAR_00, +PLEASURE_00, +LIKE_00, +DISLIKE_00, +SADNESS_00, +SORROW_00, +PRIDE_00, +HAPPY_00, +COMFORT_00, +DANGER_00, +FREE_00, +SAFETY_00, +FAILURE_00, +SAD_00, m +SAD_00, LOVE_00, +NERVOUS

   ***1b) BEHAVIOUR/EMOTION:***
   +COURAGE_00 , CRUELTY_00, +STUPID_00, pos+DAMAGE_00, +DAMAGE_00, +DAMAGE_01, +VIOLENCE_00, FRIGHTEN_00, +PROTECT_00, +FUNNY_00, +DEFENCE_00,+DANGEROUS_00, +DESTRUCTION_00, +IMPROVEMENT_00, +PUNISHMENT_00, +VIOLENCE_00 (+BATTLE_00; +FIGHT_00; +WAR_00), +COMPETITION_00.

   ***2) VERACITY:***
   +WRONG_00; n +BE_01 +LEGAL_00, +LEGAL_00.

   ***3) VITALITY:***
   +DIE_00, +, pos +DIE_00, +SICK_00, +ILL_00, +DEAD_00, +FEVER_00;+PAIN_00, +INJURY_00; +WOUND_00, +EFFORT_00, +SUFFER_00, +DEATH_00, +PROTECTION_00, +SEX_00, +HURT_00, +HARDSHIP_00.

   ***4) AESTHETICS:***
   +BEAUTY_00, +BEAUTIFUL_00.

   ***5) PROMINENCE:***
   +IMPORTANT_00, +VICTORY_00, +SERIOUS_00.

   ***6) RELIGION:***
   +GOD_00, +SAINT_00.

*7) ECONOMY:*
+MONEY_00.

*8) INTELLECT:*
+INTELLIGENCE_00, +WISDOM_00, +DIFFICULT_00.

4. Examples of axiologically-sensitive entities and their axiologically-loaded predications in meaning postulates. The case of some basic concepts included in +FEELING_00:

**#ABSTRACT>> #ATTRIBUTE >> +ATTRIBUTE_00 >> PSYCHOLOGICAL_ ATT_00 >> +FEELING_00:**

+**DESIRE_00:** +(e1: +BE_00 (x1: +DESIRE_00)Theme (x2: +FEELING_00)Referent) *(e2: +FEEL_00 (x3: (e3: +DESIRE_01 (x4: +HUMAN_00)Theme (x5)Referent)) Agent (x4)Theme (x1)Attribute)
+**PITY_00:** (f1: (e4: +BE_01 (x6)Theme (x7: +SAD_00) Attribute)Result)
**+PLEASURE_00:** +(e1: +BE_00 (x1: +PLEASURE_00) Theme (x2: +FEELING_00) Referent) *(e2: +FEEL_00 (x3: (e3: +LIKE_00 (x4)Agent (x5: +HUMAN_00) Theme) Agent (x5)Theme x1)Attribute)
**$ELATION_00:** +(e1: +BE_00 (x1: $ELATION_00)Theme (x2: +PLEASURE_00) Referent)
**+SADNESS_00:** +(e1: +BE_00 (x1: +SADNESS_00)Theme (x2: +FEELING_00) Referent) *(e2: +FEEL_00 (x3: (e3: +DISLIKE_00 (x4)Agent (x5: +HUMAN_00) Theme))Agent (x5)Theme)

5. Moreover, there are cases of non axiologically-sensitive entities, but axiologically-loaded predications in meaning postulates, as can be seen in this example:

(A) #OBJECT >> #COLLECTION >> +GROUP_00 >>+PEOPLE_00>> +POLICE_00

**+POLICE_00:** +(e1: +BE_00 (x1: +POLICE_00)Theme (x2: +ORGANIZATION_00) Referent) *(e2: +PROTECT_00 (x1)Theme (x3: +HUMAN_00)Referent) *(e3: +PROTECT_00 (x1)Theme (x4: +CORPUSCULAR_00)Referent)(e4: +HAVE_00 (x3)Theme (x4)Referent))

(B) #OBJECT >> #FEATURE >> +PART_00 >> +PLANT_PART_00 +FLOWER_00

**+FLOWER_00:** *(e1: +BE_00 (x1: +FLOWER_00)Theme (x2: +PLANT_PART_00) Referent) *(e2: n +LIVE_00 (x1)Theme (f1: +LONG_01)Duration) ***e3: +BE_01 (x1)Theme (x3: +BEAUTIFUL_00)Attribute)

6. In conclusion, the only axiological hierarchy that can be assumed at conceptual level is built into language and depends, for its relevance (positive or negative), on what is perceived by the vast majority of speakers of linguistic communities. The proposal to insert axiological notations in the FGKB ontology, in the lexica under construction or, alternatively, in other levels of meaning description in the Lexical Constructional Model (LCM) should be explored as a key factor for meaning construction.

## 5 References

Cortés-de-los-Ríos, Mª E. (2001): *Nuevas perspectivas lingüísticas en la publicidad impresa anglosajona*. Almería: University of Almería.

Coseriu, E. (1967): Lexicalische solidaritäten. In *Poetica*, 1, 293-303.

Faber, P. & Mairal Usón, R. (1999): *Constructing a lexicon of English verbs*. Berlin / New York: Mouton de Gruyter

Felices-Lago, Á. (1991): *El componente axiológico en el lenguaje. Su configuración en los adjetivos que expresan emociones y conducta en la lengua inglesa*. Ph.D. Dissertation. Granada: University of Granada.

Felices-Lago, Á. (1997): The Integration of the axiological classeme in an adjectival lexicon based on functional-lexematic principles, In C.S. Butler, et alii (eds.): *A Fund of Ideas: Recent Developments in Functional Grammar*, Amsterdam: IFOTT, 95-112.

Hare, R.M. (1952): *The language of morals.* London: Clarendon Press. (Oxford University Press, 1964).

Krzeszowski, T. P. (1990): The axiological aspect of idealized cognitive models. In Jerzy Tomaszczyk / Barbara Lewandowska (eds): *Meaning and lexicography.* Amsterdam: John Benjamins, pp. 135-165.

Krzeszowski, T.P. (1993): The axiological parameter in preconceptual image schemata. In Geiger, R.A. & Rudzka-Ostyn, B. (eds): *Conceptualizations and mental processing in language.* Berlin: Mouton de Gruyter, 307-330.

Krzeszowski, T. P. (1997): *Angels and devils in hell. Elements of axiology in semantics.* Warsaw: Energeia.

Mairal-Usón, R. & Periñán-Pascual, C. (2009): "The anatomy of the lexicon component within the framework of a conceptual knowledge base. *Revista Española de Lingüística Aplicada*, 22, 217-244.

Mairal-Usón, R. & Periñán-Pascual, C. (2010): "Role and reference grammar and ontological engineering". In J.L. Cifuentes, A. Gómez, A. Lillo, J. Mateo, F. Yus (eds.) *Los caminos de la lengua. Estudios en homenaje a Enrique Alcaraz Varó*. Alicante: University of Alicante, 649-665.

Mairal-Usón, R. & Ruiz-de-Mendoza-Ibáñez, F.J. (2008): New challenges for lexical representation within the lexical-constructional model". *Revista Canaria de Estudios Ingleses*, 57, 137-158.

Mairal-Usón, R. & Ruiz-de-Mendoza-Ibáñez, F.J. (2009): "Levels of description and explanation in meaning construction". In Butler, Ch. and Martín-Arista, J. (eds.) *Deconstructing constructions.* Ámsterdam: John Benjamins, 153–198.

Martín-Mingorance, L. (1987): "Classematics in a functional-lexematic grammar of English", *Actas del X Congreso de AEDEAN,*AEDEAN, Zaragoza, pp.377-388.

Nute, D. (2003): "Defeasible logic: theory, implementation, and applications," In Oskar Bartenstein, Ulrich Geske, Markus Hannebauer, Osamu Yoshie (eds.), *Web Knowledge Management and Decision Support:Proceedings of INAP 2001, 14th International Conference on Applications of Prolog (Revised Papers),* Berlin: Springer-Verlag, 151-169.

Osgood, Ch., Suci, G. & Tannenbaum, P. (1957): *The measurement of meaning.* Chicago: Illinois University Press.

Periñán-Pascual, C. & Arcas-Túnez, F. (2004): Meaning postulates in a lexico-conceptual knowledge base, *15th International workshop on databases and expert systems applications,* IEEE, Los Alamitos (California), pp. 38-42.

Periñán-Pascual, C. & Arcas-Túnez, F. (2005): Microconceptual-knowledge spreading in FunGramKB. *Proceedings on the 9th IASTED international conference on artificial intelligence and soft computing*, ACTA Press, Anaheim-Calgary-Zurich, pp. 239-244.

Periñán-Pascual, C. and Arcas-Túnez, F. (2007): "Cognitive modules of an NLP knowledge base for language understanding. *Procesamiento del lenguaje natural* 39, 197-204.

Periñán-Pascual, C. & Mairal-Usón, R. (2009a): Bringing Role and Reference Grammar to natural language understanding. *Procesamiento del Lenguaje Natural*, 43, 265-273.

Periñán-Pascual, C. & Mairal-Usón, R. (2009b): Formalizing conceptual knowledge. [Unpublished text].

Pottier, B. (1964): Vers une sémantique moderne. *Travaux de linguistique et literatura* 2.1, 101-137.

Ruiz-de-Mendoza-Ibáñez, F. J. & Mairal-Usón, R. (2008): Levels of description and constraining factors in meaning construction: an introduction to the Lexical Constructional Model. *Folia Linguistica* 42/2, 355-400.

Velardi, P., Pazienza, M.T. & Fasolo, M. (1991): How to encode semantic knowledge: a method for meaning representation and computer-aided acquisition. *Computational Linguistics* 17, 2. 153-170.

***

# Book Review


**Mündliche Fachtexte der französischen Rechtssprache.**

Julia Neu

Frank & Timme: Berlin 2011. 294 Seiten.
Forum für Fachsprachenforschung, Band 96.

ISBN: 978-3-86596-351-2.

www.frank-timme.de

**Rezensiert von**:
Anne Lise Kjær
Juristische Fakultät, Universität Kopenhagen
Kopenhagen, Dänemark
alk@jur.ku.dk

## Wie fachlich sind eigentlich Rechtstexte? Buchbesprechung

### 1 Thema und Forschungstradition des Buches

Wie der Titel angibt, sind Thema des Buches - der Promotionsarbeit der Autorin - mündliche Fachtexte der französischen Rechtssprache. Im Mittelpunkt der Arbeit steht die doppelte Frage „wie sich die Faktoren „Mündlichkeit" und „Fachlichkeit" in Texten manifestieren und in welcher Beziehung sie dabei zueinander stehen" (S.14). Dieser Frage geht Julia Neu nach, indem sie mündliche Kommunikationssituationen im Bereich des französischen Notariatswesens einer systematischen Analyse unterzieht.

Insbesondere die Gründlichkeit und Sorgfalt der Analysen und die Stringenz des selbständig aufgebauten Analysemodells sind bei der Untersuchung hervorzuheben. Die theoretische Grundlage der Arbeit ist dagegen weniger überzeugend; zu beanstanden ist vor allem, dass Julia Neu im Buch nicht begründet, warum sie gerade

Rechtstexte als empirische Datengrundlage für ihre Untersuchung zur Fachlichkeit mündlicher Texte gewählt hat. Die Besonderheiten von Rechtstexten - ihre starke Bindung an einem Rechtssystem und dessen Institutionen – machen sie doch weniger geeignet als Gegenstand einer Analyse, die allgemeine Schlussfolgerungen über mündliche Fachtexte ziehen will.

Entsprechend dem Ziel der Untersuchung bildet die allgemeine Fachsprachenlinguistik den theoretischen Rahmen der Untersuchung, während ein rechtslinguistischer Blickwinkel nur am Rande der Untersuchung angelegt wird. Diese theoretische Wahl hängt vermutlich mit dem akademischen Hintergrund der Verfasserin zusammen.

Julia Neu studierte am Fachbereich Angewandte Sprach- und Kulturwissenschaft an der Johannes Gutenberg-Universität Mainz in Germersheim und ist seit 2004 in Germersheim am Arbeitsbereich Interkulturelle Germanistik als Dozentin tätig. Ihre starke Verankerung in der deutschen Fachsprachenlinguistik lässt sich an der Beschränkung auf nur in der deutschen angewandten Sprachwissenschaft bekannte Forschungstraditionen deutlich erkennen. Die Autorin beruft sich weitgehend auf führende deutsche bzw. deutschsprachige Fachsprachenforscher, die seit Jahrzehnten im Bereich der Fachtextlinguistik tätig sind und waren: Kirsten Adamzik, Klaus-Dieter Baumann, Hartwig Kalverkämper, Hans-Rüdiger Fluck, Rosemarie Gläser, Susanne Göpferich, Lothar Hoffmann, Dieter Möhn, Els Oksaar, Heribert Picht, Wolfgang Pöckl, Hartmut Schröder u.a.. Auch wenn andere bzw. neuere Ansätze einbezogen werden, bildet die traditionelle deutsche Fachsprachenforschung den Ankerpunkt der Arbeit, die also als „Normalwissenschaft“ im Sinne von Kuhn bezeichnet werden kann.

Dies ist ein durchaus legitimer Rahmen der Untersuchung, die an diesem Ausgangspunkt auch zu bewerten ist. Aber der Forschungszugang schränkt die Leserschaft wesentlich ein. Aus welchen Gründen, für welche Zwecke und für wen sind die in der Arbeit untersuchten Fragen interessant und relevant? Ich nehme an, dass nur andere (deutsche) Fachsprachenlinguisten ohne weiteres anerkennen, dass das Buch eine Forschungslücke füllt, die es wert ist auszuforschen. Andere Leser sind darauf angewiesen, die Stringenz und Gründlichkeit der Analysen zu bewundern.

## 2 Aufbau, Inhalt und Gegenstand der Untersuchung

Was also will die Autorin? Sie will eine in der Fachsprachenlinguistik mehrmals hervorgehobene Lücke füllen: sie will fachbezogene mündliche Kommunikation als Fachtextsorte systematisch beschreiben und so einer von Rosemarie Gläser schon 1985 formulierten Aufforderung nachkommen.[1] Dabei lehnt sie es ausdrücklich ab, eine auf den interaktionalen Charakter mündlicher Texte ausgerichtete gesprächsorientierte Analyse durchzuführen. Sie will vielmehr Gespräche als Fachtexte studieren und der Frage nachgehen, einerseits wie die Fachlichkeit mündlicher Texte zum Ausdruck

[1] Rosemarie Gläser (1985, 16).

kommt, und andererseits wie die Mündlichkeit der Fachkommunikation die Texte prägt.

Zu diesem Zweck stellt sie ein detailliertes Analysemodell auf, das sie auf das erhobene Datenmaterial empirisch testet: sieben Stunden Aufnahmen von Kommunikationssituationen im Bereich des französischen Notariatswesens sowie die entsprechenden Transkriptionen.

Die Untersuchung besteht aus 3 Teilen: (1) der theoretischen Auseinandersetzung mit den Begriffen „Text“, „Mündlichkeit“ und „Fachkommunikation“ (Kapitel 2-4; insgesamt 64 Seiten); (2) der Ausarbeitung des Analysemodells (Kapitel 5-6; insgesamt 65 Seiten); und (3) der Analyse des Korpus (Kapitel 7-9; insgesamt 120 Seiten). Dazu kommen eine kurze Einleitung (Kapitel 1; 5 Seiten), eine ebenso kurze Zusammenfassung und die Schlussgedanken der Autorin (Kapitel 10-11; insgesamt 6 Seiten) außer einem Literaturverzeichnis (22 Seiten).

Wie aus dem Inhaltsverzeichnis ersichtlich ist, macht die empirische Analyse den Schwerpunkt der Untersuchung aus, während die theoretischen Ausführungen und Schlussfolgerungen eher kurzgefasst sind. Im folgenden werde ich mich aber gerade mit den theoretischen Annahmen der Autorin auseinandersetzen und ihren m.E. zu engen Blick auf Kommunikation in einem institutionellen Rahmen des Rechts beanstanden.

**3 Fachlichkeit bzw. Rechtlichkeit: Theoretische Fehlschlüsse?**

Wie oben beschrieben will die Autorin die Mündlichkeit fachlicher Texte bzw. die Fachlichkeit mündlicher Texte untersuchen, einschließlich der gegenseitigen Beziehung dieser beiden Faktoren. Nirgends begründet sie aber die Wahl der untersuchten Textsorte: mündliche Texte im Notarwesen. Aus dem Blickwinkel der Rechtslinguistik scheinen Rechtstexte als eine schlechte Wahl, wenn allgemeine Schlussfolgerungen hinsichtlich des Untersuchungsgegenstandes gezogen werden sollen. Nicht alle Rechtstexte sind ohne weiteres als bloße Fachtexte einzustufen – und selbst wenn Rechtstexte auch Merkmale von Fachlichkeit aufweisen, sind sie vor allem durch ihre institutionelle Einbindung in einer Rechtsordnung und ihren Handlungscharakter gekennzeichnet.

Julia Neu setzt sich im Buch zwar mit diesem Problemkreis auseinander (Seite 60-65); die Gründlichkeit ihrer Überlegungen lässt aber einiges zu wünschen übrig.

Ausgehend einerseits von Kalverkämpers Beschreibung von „Fachlichkeit“ (Kalverkämper 1996, 132) und andererseits von der Definition des Begriffs „Recht“ im Creifelds Rechtswörterbuch (2002) will die Verfasserin Rechtstexte nur als Fachtexte auffassen, wenn sie auf die Rechtsordnung bezogenes Handelen manifestieren (S. 60). Eine Handlungsperspektive auf Rechtstexte anzulegen ist durchaus sinnvoll; sie als Kennzeichen von der *Fachlichkeit* von Rechtstexten zu bestimmen scheint aber geradezu falsch. Ganz im Gegenteil ist das rechtsordnungbezogene Handeln durch Rechtstexte wohl Kennzeichen ihrer *Rechtlichkeit*, also gerade *das* Kennzeichen von Rechtstexten, das sie von anderen (Fach)texten abhebt.

Insbesondere wenn sich die Autorin auf Kalverkämper und dessen Verständnis von „Fachlichkeit" beruft, überzeugt ihre Definition von Rechtstexten nicht: „Kalverkämper verdeutlicht den Zusammenhang von Kommunikation und Fachlichkeit am Beispiel eines Apfels, der nicht per se ein fachlicher Gegenstand ist, aber zu einem solchen wird, wenn man als Biologe, Theologe, Lehrer oder Mediziner über ihn spricht" (S. 60). Eine solche Auffassung von Fachlichkeit hat wenig zu tun mit dem Handlungscharakter der Rechtskommunikation. Die Übertragung von Kalverkämpers Fachlichkeitsbegriff auf Rechtstexte führt deshalb leicht zu Fehlschlüssen: Recht wird erst zu Recht, wenn Juristen darüber sprechen?

Nein, meint Julia Neu selbst. Ganz im Gegenteil: Ausgebildete Juristen müssten nicht notwendig an einem Kommunikationsvorgang beteiligt werden, um ihn als Rechtskommunikation zu bestimmen. Auch Diskussionen zwischen nicht-Juristen bzw. Laien könnten als Rechtskommunikation gelten, wenn sie sich bloß auf die Rechtsordnung beziehen. Dieser Auffassung stimme ich zu; sie steht aber im scheren Kontrast zu dem von Kalverkämper vorgeschlagenen Fachlichkeitsbegriff, auf den sich die Verfasserin auch beruft.

Julia Neus Begründung, die in der deutschen Rechtslinguistik allgemein anerkannten Typologisierungen von Rechtstexten (z.B. Engberg 1993, Busse 2000 und Wiesmann 2004) für ihren Untersuchungszweck abzuweisen, genügt nicht: „Während es dort notwendig ist, das Gebiet der Rechtstexte mit Hinblick auf eine bestimmte Fragestellung einzugrenzen, soll das von mir entwickelte Analysemodell sich auf möglichst viele Texte dominant mündlicher Fachkommunikation anwenden lassen, die einen unterschiedlichen Grad an Prototypizität aufweisen." (S. 64). In dem Sinne liegt „[e]ine Diskussion zwischen einer Kassiererin und einer minderjährigen Kundin über den Kauf von Alkohol [...] hierbei in der Peripherie der ‚Rechtskommunikation', wohingegen eine Gerichtsverhandlung im Zentrum befindet" (S 63).

Was Neu bei ihrer Prototypizitätsperspektive zu vergessen scheint, sind Texte der fachinternen juristischen Kommunikation, wie z.B. Vorträge juristischer Professoren über rechtliche Themen oder Diskussionen zwischen Lehrer und Studenten der Rechtswissenschaft während der Juristenausbildung: sind solche Texte mehr oder weniger prototypische Beispiele von Rechtstexten als z.B. alltägliche Rechtskommunikationen im Detailwarenhandel? Neu kommentiert dieses Dilemma nicht und bestimmt wahrscheinlich alle Texte, die sich in irgendeiner Weise auf eine Rechtsordnung beziehen, als Rechtstexte, ohne den Maßstab anzugeben, an dem den Grad rechtstextlicher Prototypizität gemessen wird.

Neu erklärt nicht, warum ihre Fragestellung eine weitere Definition von Rechtstext erfordert als die sonst üblichen, und es bleibt unklar, warum sie sich nicht an den handlungsbezogenen Definitionen von Rechtstexten orientiert, die die von ihr genannten Forscher vertreten.

## 4 Das Korpus bzw. die Korpora

Ich bin der Auffassung, dass Julia Neu zwischen Texten der *Juristenkommunikation* und Texten der *Rechtskommunikation* hätte unterscheiden sollen, und ihre Daten entsprechend in fachliche (informative) bzw. rechtshandlungsvollziehende (performative) Texte hätte einteilen sollen. „Fachlich“ in der von Kalverkämper vorgeschlagenen Bedeutung sind nur Texte der Juristenkommunikation, während Texte der Rechtskommunikation nicht notwendigerweise Fachlichkeit aufweisen, wie Neu zu Recht annimmt. Texte der Rechtskommunikation kommen u.U. ohne die zur Rechtswissenschaft gehörende Fachsprache aus; juristische Fachsprache ist zur Durchführung von Rechtshandlungen nicht in allen Fällen notwendig.

Neu teilt ihr Material tatsächlich in zwei Teilkorpora ein: Teilkorpus 1 „enthält Aufnahmen, die mir ein Notariat in Elsass ermöglichte“ (S. 131) und umfasst die Textsorten ‚Abschluss und Beurkundung eines Vertrags’, ‚notarielles Beratungsgespräch’, ‚Arbeitsbesprechung mit einem Praktikanten’ und ‚Errichtung eines Inventars’. Teilkorpus 2 besteht aus mehreren Vertretern der Textsorte ‚Übung eines Prüfungsgesprächs’, also Texten im Bereich der Notarausbildung, die bei einem Besuch an einer Universität behoben wurden. Die von der Verfasserin vorgenommene Einteilung beruht also auf den Ort des Zustandekommens der Texte, nicht aber auf ihren Grad an Fachlichkeit bzw. Handlungsbezogenheit. Die Begründung dafür leuchtet mir nicht ein.

Eine alternative Einteilung des Korpus in informative bzw. performative Rechtstexte hätte die Teilanalysen der Untersuchung nicht nur aus dem Blickwinkel der Rechtslinguistik interessanter gemacht, sondern auch die von der Verfasserin erwünschte Brauchbarkeit der Untersuchungsergebnisse für die Fachtextlinguistik im allgemeinen verbessert. Rechtstexte als bloße Beispiele für Fachtexte zu behandeln, ohne ihre Besonderheit – ihre Institutionalität und Bindung an zum Teil gesetzlich vorbestimmten Rechtshandlungen einer bestimmten Rechtsordnung – hervorzuheben und im Untersuchungsdesign zu berücksichtigen, bleibt meines Erachtens nicht sinnvoll.

## 5 Zum Schluss

Wie aus meinen obigen Kommentaren hervorgeht, finde ich das Buch von Julia Neu für eine Leserschaft relevant, die sich wie Neu selbst im Zentrum der deutschen Fachsprachenlinguistik befindet. Für andere fehlen der Arbeit Reflexionen über die Brauchbarkeit der Untersuchungsergebnisse: Warum sollten wir uns dafür interessieren? Was ich als Rechtslinguistin besonders beanstande, ist die Behandlung von Rechtstexten als eher zufällige Beispiele von Fachtexten.

Mit einer Erweiterung der Perspektive kann das Untersuchungsmaterial in einem theoretischen Nachwort der Verfasserin auf Fragen der Verständlichkeit, Legitimität und Autorität des Rechts und des Juristenstandes überprüft werden:

- Verständlichkeit von Rechtstexten: sind mündliche Rechtstexte verständlicher als schriftliche?
- Nähekommunikation und Bürgernähe der Justiz und Verwaltung: wie verhalten sie sich zueinander?
- Ist die Macht der Juristen auch eine sprachliche Macht: wenn die Juristen ohne Fachsprache in mündlichen Kommunikationssituationen auskommen, warum sind dann Fachtermini und komplexe Syntax konstante Parameter schriftlicher Rechtstexten?
- Usw.

Die Daten sind erhoben – neue Analysen zur *Rechtlichkeit* der untersuchten Kommunikationssituationen sind willkommen.

**Literaturhinweise**

Busse, Dietrich (2000): Textsorten im Beriech des Rechtswesens und der Justiz. In Brinker, Klaus et al. (Hrsg.): *Text- und Gesprächslinguistik. Ein internationales Handbuch zeitgenössischer Forschung.* (S. 658-675) Berlin: de Gruyter 2000 (= Handbücher der Sprach- und Kommunikationswissenschaft, 16).

Creifelds, Carl ([17]2002): *Rechtswörterbuch.* München: Beck [1968]

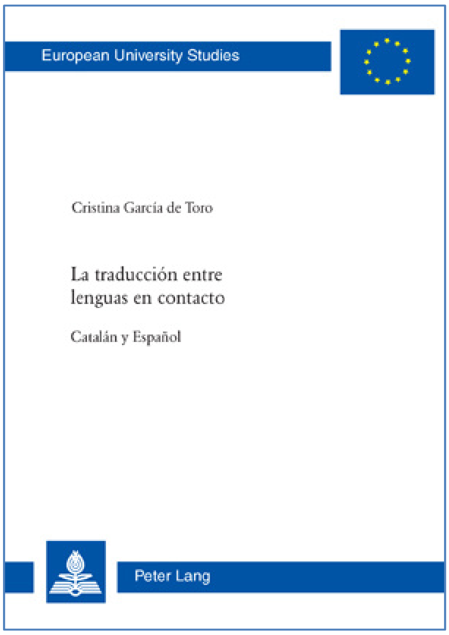
Engberg, Jan (1993): Prinzipien einer Typologisierung juristischer Texte. Fachsprache. Internationale Zeitschrift für Fachsprachenforschung, -didaktik und Terminologie, 15 (1-2): 31-38.

Gläser, Rosemarie (1985): Standortbestimmung einer Fachtextlinguistik. In Gläser, Rosemarie (Hrsg.): *Fachsprachliche Textlinguistik. Vorträge der sprachwissenschaftlichen Arbeitstagung an der Sektion Fremdsprachen der Karl-Marx-Universität Leipzig am 11. und 12. Dezember 1984* (S. 2-19). Berlin: Akademie der Wissenschaften der DDR. Zentralinstitut für Sprachwissenschaft (= Linguistische Studien. Reihe A. Arbeitsberichte, 133).

Kalverkämper, Hartwig (1996): Im Zentrum der Interessen: Fachkommunikation als Leitgröße. *Hermes. Journal of Linguistics*, 16: 117-176.

Wiesmann, Eva (2004): *Rechtsübersetzung und Hilfsmittel zur Translation. Wissenschaftliche Grundlagen und computergestützte Umsetzung eines lexikographischen Konzepts*. Tübingen: Narr (=Forum für Fachsprachenforschung, 65).

***




**La traducción entre lenguas en contacto. Catalán y español.**

Cristina García de Toro

Peter Lang. Bern 2009. 182 p.
European University Studies

ISBN: 978-3-03911-493-1

www.peterlang.net

**Autora de la reseña:**
Ona Domènech Bagaria
Estudis d'Arts i Humanitats, Universitat Oberta de Catalunya, Barcelona, Spain
odomenechb@uoc.edu

El libro de Cristina García de Toro nos ofrece una panorámica de la traducción entre parejas de lenguas que se encuentran territorialmente en contacto, tomando como base la pareja formada por el español y el catalán. La originalidad de la obra radica precisamente en su tema de estudio, puesto que la traducción entre lenguas en contacto no ha suscitado demasiado interés entre los traductólogos, seguramente porque, como afirma García de Toro, existe la asunción generalizada de que los hablantes de esas zonas son bilingües, de modo que la traducción no es necesaria, sobretodo si se trata de lenguas muy cercanas, como es el caso del español y el catalán. La realidad, sin embargo, nos demuestra que la traducción entre lenguas que conviven territorialmente es muy abundante, aunque no siempre persigue el objetivo de la inteligibilidad; y que las personas que conviven en estos territorios no suelen ser bilingües, puesto que siempre hay algún tipo de diglosia, fruto de la tensión política e ideológica existente entre las lenguas en contacto.

El libro parte de trabajos previos realizados por la autora, algunos inéditos y otros ya publicados. Y posiblemente esta condición recopilatoria explica la diversidad de enfoques que presentan los distintos temas tratados: desde la perspectiva más teórica de los primeros capítulos, pasando por los enfoques más descriptivos de los capítulos centrales hasta llegar a la perspectiva más aplicada de los capítulos que cierran el libro. Cabe mencionar, sin embargo, un meritorio hilo conductor presente a lo largo de toda la obra: el de destacar, de entre los temas tratados, los que presentan interés como posibles líneas de investigación, con la voluntad de invitar al estudio de la traducción entre lenguas que comparten un mismo territorio.

La primera parte del libro, integrada por los tres primeros capítulos, presenta una perspectiva teórica del tema, que se inicia con la voluntad de hallar la denominación más acertada para referirse a la traducción entre las lenguas que están en contacto dentro de un determinado territorio, como es el caso del catalán y el castellano en los territorios catalanohablantes del Estado español. Para ello, en el capítulo primero García de Toro se centra en la valoración de las denominaciones utilizadas por las disciplinas con más tradición en el estudio de las lenguas en contacto, la sociolingüística y la etnolingüística, que manejan esencialmente dos opciones: *traducción entre comunidades de habla bilingües* y *traducción entre lenguas en contacto*. Después de revisar las nociones de *comunidad de habla*, *bilingüismo* y *lenguas en contacto*, la autora concluye que esta última es la que presenta mayores ventajas, de modo que considera más operativa la designación *traducción entre comunidades con lenguas en contacto* para referirse a situaciones como las del territorio español.

El segundo capítulo consiste en un breve recorrido por la situación lingüística de la Unión Europea, con el objetivo de ubicar y describir la situación actual de las lenguas en España a partir del contraste con las situaciones vecinas. La autora parte de un representativo trabajo de Siguan titulado *La Europa de las lenguas*, para explicar las diferentes políticas lingüísticas que se siguen actualmente en Europa. Y finaliza concluyendo que la situación de *autonomía lingüística* propia de las lenguas del Estado español constituye una posición intermedia en la escala que presenta dicho autor, en uno de cuyos extremos estarían las situaciones menos favorables a la pluralidad lingüística (*monolingüismo* y *protección a las minorías*) y en el otro, aquellas en las que se reconocen mayores grados de igualdad entre las lenguas (*federalismo lingüístico* y *plurilingüismo institucional*).

A continuación, partiendo de datos concretos sobre el conocimiento y uso de las distintas lenguas en España, García de Toro nos muestra la situación sociolingüística entre las diversas comunidades, que es muy distinta, en función, básicamente, del nivel de competencia lingüística. Así, zonas como Cataluña, Baleares y Galicia, con un nivel alto de competencia (buena parte de la población se declara bilingüe), tienen un porcentaje elevado de uso de la lengua propia. En cambio, los porcentajes de uso son mucho menores en las comunidades de lengua euskera, donde la proporción de población que entiende dicha lengua no llega a la mitad. La misma relación se da en las comunidades de habla catalana: los índices más altos de competencia y uso del catalán se dan en Cataluña y en las islas Baleares, mientras que los niveles son bastante más bajos en la Comunidad Valenciana. Finalmente, la autora acaba este recorrido con una breve e interesante explicación de tres situaciones de contacto que se dan en Europa, más allá de las fronteras del Estado español, en las que una de las lenguas es el catalán. Se trata de los casos de Andorra, la ciudad de Alguer (en la isla de Cerdeña, Italia) y el Rosellón o Cataluña Norte (en Francia).

Esta primera parte preliminar del libro se cierra con un tercer capítulo que pretende ofrecer una denominación, satisfactoria desde la perspectiva traductológica, para la traducción en las diferentes situaciones de convivencia de lenguas. García de Toro muestra, muy acertadamente, que todas las denominaciones posibles presentan algún aspecto polémico, puesto que incluyen términos problemáticos como *lenguas maternas*, *primeras lenguas*, *lenguas extranjeras*, *lenguas A*, *etc.* Y finalmente defiende la denominación *traducción entre lenguas A*, ya que a pesar de ser la que goza de menor tradición en la investigación traductológica, es la única exenta de los componentes ideológicos y las cargas valorativas asociadas al resto de las opciones posibles.

La segunda parte del libro está compuesta por tres capítulos –del cuarto al sexto–, que ahondan en la caracterización de la traducción entre el catalán y el español. El capítulo cuatro esboza las características generales de dicho proceso en las diversas zonas catalanohablantes del Estado español, destacando el hecho de que se trata de una traducción que, mayoritariamente, no responde al criterio de inteligibilidad, sino que va ligada al proceso de normalización del catalán. Esta perspectiva fundamental, que visibiliza la importancia de la traducción como instrumento de normalización lingüística, ya había sido destacada y explicada anteriormente por otros autores, tal y como menciona la propia García de Toro; por ello consideramos que la aportación más relevante de este capítulo consiste en el hecho de que, además de estos condicionantes políticos, se enumeran y explican otros factores importantes que intervienen en la traducción catalán-español y la caracterizan, a saber: la convivencia de las lenguas, la proximidad formal, la especificidad temática del mercado laboral y la especificidad de la enseñanza universitaria de la traducción en esta combinación lingüística.

Este último punto es retomado en el capítulo siguiente –el quinto–, que se adentra en las características de la hibridación que presentan los grupos de alumnos en el aula de traducción entre catalán y español, que responde al hecho de que, desde el punto de vista traductológico, cuando dos lenguas conviven ninguna de las dos se considera como lengua extranjera, de modo que al traductor profesional se le supone la competencia bidireccional. Si a ello le sumamos la no distinción entre la lengua materna de los traductores, el resultado son los contornos difusos entre la *traducción directa* y la *traducción inversa* que se dan en esta práctica de traducción. También se aborda el fenómeno de la *autotraducción*, estrechamente relacionado con la bidireccionalidad propia de la traducción entre catalán y español, puesto que un buen número de autores literarios, con un elevado grado de bilingüismo, suelen traducir sus propias obras, de modo que las fronteras entre traducción y versión se diluyen. La autora considera que el binomio *traducción entre lenguas en contacto* y *autotraducción* abren un camino de investigación muy atractivo en el panorama traductológico actual.

El capítulo 6 cierra la segunda parte del libro dibujando las perspectiva profesionales y de investigación que ofrece la traducción entre catalán y español. Desde el punto de vista profesional, el ámbito administrativo es el que muestra una situación más normalizada, mientras que en el resto de ámbitos (audiovisual, informático, publicitario, científico-técnico) la traducción entre esta pareja de lenguas todavía es bastante minoritaria. También la investigación entre catalán y español es incipiente: aunque existen buenos estudios históricos de traducciones literarias y numerosos estudios sobre la traducción administrativa, es urgente, según García de Toro, ampliar el panorama investigador a otros ámbitos como la traducción audiovisual y de productos informáticos, el proceso cognitivo de la traducción entre estas lenguas, el grado de bilingüismo de la traducción, la didáctica de la traducción o la crítica y evaluación de traducciones.

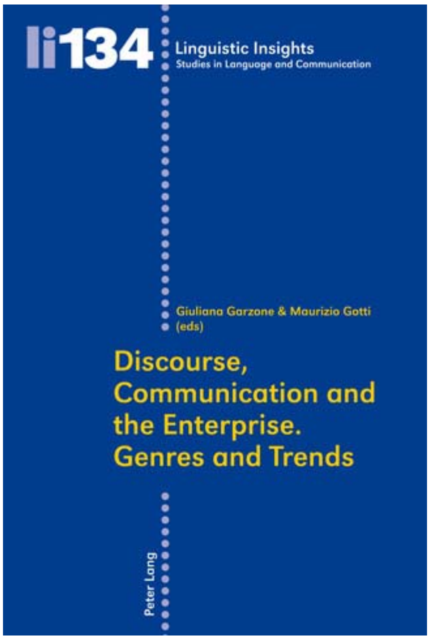
La tercera parte del libro –capítulos siete y ocho–, está dedicada al ámbito que goza de más consolidación profesional en la traducción del español al catalán: la traducción administrativa. En el capítulo siete la autora explica las características principales de este tipo de traducción, a partir de las obras clásicas de diversos autores, entre los cuales destaca Carles Duarte, quien ha tratado magistralmente los temas esenciales relacionados con el lenguaje jurídico y administrativo catalán. En este completo y fundamentado retrato de la traducción administrativa del catalán al español, sólo echamos de menos algunas distinciones terminológicas, que a nuestro parecer son importantes, como la delimitación entre los adjetivos “administrativo” y “jurídico”, aplicados tanto a la traducción, como al lenguaje o a los textos que se generan.

Esta tercera parte se cierra con el capítulo ocho, en el cual, después de exponer diversas clasificaciones sobre la documentación administrativa, la autora presenta una propuesta de clasificación realizada en el marco de un proyecto de investigación sobre géneros textuales para la traducción. Se trata de una clasificación amplia, que incluye un gran abanico de géneros, con la finalidad de que puedan ser usados como textos paralelos, modelos de documentos, etc. La originalidad y acierto de dicha tipología radica en el hecho de que no se dirige a ningún organismo concreto, ni tampoco al ciudadano en general, sino que está pensada por y para traductores, lo que la lleva a ser un instrumento muy útil al servicio de la traducción y la redacción administrativa.

Y llegamos así a la cuarta y última parte del libro, dedicada a la traducción literaria del catalán al español, uno de los ámbitos más consolidados en esta dirección de traducción entre las dos lenguas. Esta parte del libro es la más extensa –puesto que ocupa los capítulos nueve, diez, once y doce– y también la menos generalista, en el sentido de que se centra en el análisis de la traducción del catalán al español de tres novelas concretas: *No emprenyeu el comissari* i *Penja els guants Butxana*, del autor valenciano Ferran Torrent, i *Vida de gos i altres claus de vidre*, del autor catalán Jaume Fuster. Tanto la extensión de esta parte, como la profundidad y el detalle del análisis que se presenta se explican por el hecho de que la autora recoge aquí las conclusiones de su tesis doctoral, tal y como ella misma menciona al inicio del capítulo nueve. El detalle del estudio llevado a cabo proporciona un buen ejemplo de rigurosa metodología aplicable al análisis de las traducciones literarias entre catalán y español, y también nos permite comprobar hasta que punto la convivencia de las lenguas influye en las decisiones de traducción que se adoptan.

En definitiva, el libro de García de Toro nos muestra la traducción entre lenguas en contacto que conviven en un mismo territorio como una actividad singular y necesaria, que genera interesantes interrogantes, a algunos de los cuales la autora dedica los distintos capítulos de su obra: desde consideraciones de corte más teórico sobre cuál es la denominación más adecuada para referirse a este tipo de traducción, pasando por cuáles son sus características más importantes, o cuáles son los ámbitos donde se practica, y dentro de ellos, qué peculiaridades presenta. Quedan, evidentemente, otros interrogantes por responder, muchos de los cuales son mencionados por la autora en diversos momentos de la obra, que invita constantemente al estudio de dichos temas, con una actitud positiva y animada que constituye uno de los principales valores de libro. Si a ello le sumamos la singularidad del tema que se aborda, podemos concluir que la lectura de la monografía reseñada nos parece absolutamente útil y recomendable para los profesionales o estudiosos de la traducción interesados en las lenguas que conviven dentro de un mismo territorio.

***




**Discourse, communication and the entreprise. Genres and trends.**

Giuliana Garzone and Maurizio Gotti (eds)

Peter Lang. Bern 2011. 451 p.
Linguistic Insights - Volume 134

ISBN: 978-3-0343-0620-1

www.peterlang.com

**Reviewed by:**
Anne Marie Bülow
Department of International Business Communication
Copenhagen Business School, Denmark
amb.ibc@cbs.dk

This volume in the Peter Lang series *Linguistic Insights* contains conference papers from 2009 dealing with four main areas within the field of discourse and communication studies relating to business or organizational communication, viz. a section surveying and projecting hopes for the academic field, one on organizational language issues, one on corporate reporting, and finally on public relations in the sense of media and promotion studies. The majority of the contributors come from Italy, with Belgium running second.

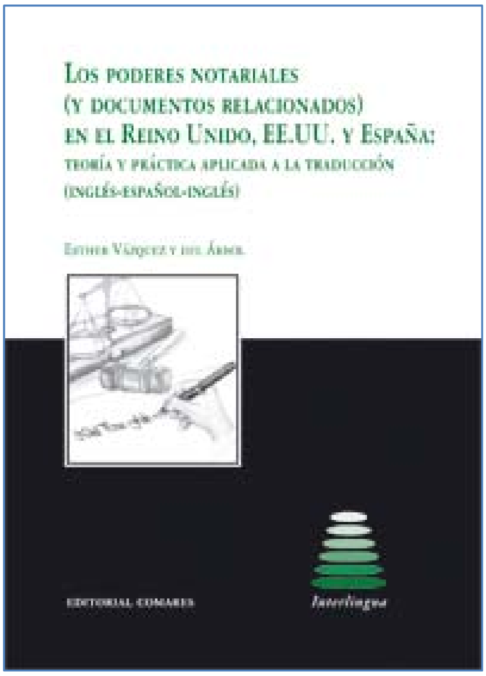
The collection shows the tension in the field between discourse studies for their own sake (for example genre studies in texts containing visions, missions and CSR and environmental reporting, which are readily available on the internet), and those that seek to relate their textual findings to the world around them. Thus Charles advocates studies of the kind practiced in Helsinki, where the researchers follow the paths of communication used by their subjects without ever considering the original nationality of the networking employees, while other contributions, more traditionally, pay attention to the specific localisation that is used in translations or parallel websites, or compare closely similar companies.

Many readers will find most interest in the studies of texts that are less familiar to Western analysts. For oral discourse, subtle power and control issues transpire in the chapter by Engelhardt, who followed meetings in a German subsidiary in the Czech Republic. Here language policy dictates that meetings should be in the language of the majority, and German expatriates therefore have to rely on the interpreter. However, when this person is not present to police the proceedings, people manage, triangulating between Czech, German and English, and the patterns of code-switching are revealing. The section on corporate reports, in its turn, contains studies that consider the ideological context. Several studies analyse annual reports from Chinese and Russian companies, and show thought-provoking instances of intertext between official state policy or legislation and the companies' efforts to construct a legitimate identity.

Genre analysis is a staple of studies in organizational discourse, and it can be a little schematic where the analyst's main contribution is ticking off points and introducing slight changes to established paradigm. But in the media section, genre analysis is put to good use with studies that follow a diachronic development. Thus Degano argues that financial news has moved from the special interest pages to headline news that is basically political; thus 'the market' is ascribed a different kind of agency.

Discursive insights are always interesting when applied to corporate or organizational texts. But with the heightened awareness of stakeholder management and multimedia affordances, it would be true progress if more analysts were numerate and took on the *full* annual reports, and not just the prose from the Communications department like the CEO's letter. If the conference's keynote spoke up for studies of real-time communication in context, it could be supplemented with a plea for more studies of texts checked against their economic or ideological context and, not least, their effects. Perhaps the next conference in the series will produce a volume that takes on that challenge.

***




**Los poderes notariales (y documentos relacionados) en el Reino Unido, EE.UU. y España : teoría y práctica aplicada a la traducción (inglés-español-inglés).**

Esther Vázquez y del Árbol

Editorial Comares. Granada 2009 345 p..

ISBN: 978-84-9836-602-0

www.comares.com

**Autora de la reseña:**
Begoña Montero Fleta
Dep. de Lingüística Aplicada, Escuela Técnica Superior Ingeniería Informática
Universidad Politécnica de Valencia, Spain
bmontero@idm.upv.es

El presente volumen constituye una importante aportación a varios campos y se caracteriza por un enfoque didáctico, fruto de años de investigación, docencia y ejercicio profesional de la autora. El libro se ha subvencionado gracias a la aportación del Grupo de Investigación HUM-383 (Interlingüística), de la Junta de Andalucía, al que pertenece la Doctora Vázquez desde hace más de una década. El prólogo ha sido elaborado por el Dr. Emilio Ortega Arjonilla, profesor titular de Universidad y coordinador de la Colección Interlingua, a la que pertenece el presente libro.

La investigación se ha organizado en base a cuatro pilares (bloques) fundamentales, que dibujan la macroestructura del manuscrito. Tras una introducción, destinada a definir y delimitar el objeto de estudio y adscribirlo a un área del derecho, en el primer bloque de capítulos (1-6) se analizan los poderes notariales en los sistemas de derecho británico, norteamericano y español. Los países cuyos documentos notariales ha investigado la autora son los que siguen: Inglaterra, Gales, Escocia, Irlanda del Norte, EE.UU. y España, por lo que el enfoque es amplio y muy variado.

Así, en el primer capítulo se relatan los rasgos típicos de los poderes notariales de Inglaterra y Gales: modalidades del documento, condiciones para poder otorgar un poder de uno u otro tipo y cuestiones relativas a la normativa pertinente.

En el segundo capítulo se aportan las cuestiones relativas a Escocia e Irlanda del Norte: tipos de poderes notariales, requisitos de otorgamiento, aspectos referidos a Escocia y los relacionados con Irlanda del Norte, así como las normas legales que rigen ambos sistemas.

En el tercer capítulo se ahonda en el ámbito norteamericano: modalidades del documento, condiciones para poder otorgar un poder notarial, legislación en vigor, y, por último, aspectos divergentes según cada Estado.

En el cuarto capítulo se describen los rasgos de la declaración de voluntades anticipadas (legislación vigente, requisitos del documento…) y sus divergencias en base al país en el que se otorgue: Inglaterra, Gales, Escocia, Irlanda del Norte, EE.UU. o España.

En el quinto capítulo se investigan los poderes notariales en España: tipología, condicionantes para su otorgamiento y legislación aplicable.

En el sexto capítulo se recapitula la información expuesta en los cinco anteriores para luego proceder a contrastar las divergencias más significativas de los sistemas de derecho y de los géneros jurídicos a estudio: denominación del poderdante y apoderado, legislación aplicable, rasgos del lenguaje empleado en los poderes, número de apoderados por poder, características del acto de otorgamiento, redacción del poder, sistema de otorgamiento y variedades de poder notarial.

El segundo bloque (capítulos 7-8) ofrece dos listados alfabéticos de elementos léxicos y fraseológicos extraídos del corpus textual analizado. De este modo, el capítulo 7 muestra un glosario inglés-español y el capítulo 8 ofrece otro español-inglés. Estos glosarios incluyen explicaciones de los problemas que pueden surgir al abordar la traducción de conceptos que no encuentran equivalente exacto en la cultura meta.

El tercer bloque viene representado por el capítulo 9, dedicado a la traducción (inglés-español) de diversos poderes notariales, declaraciones juradas, declaraciones de voluntades anticipadas y apostillas, todos ellos géneros británicos y norteamericanos representativos del derecho de representación civil y penal. Con este capítulo se ofrece al profesional y al discente una visión especialmente pragmática del documento notarial y de la trasmisión de su contenido a otra lengua y cultura.

El cuarto bloque consta de un único capítulo en el que la autora aporta una amplísima gama de formularios de poderes válidos en España, Inglaterra, Gales, Escocia, Irlanda del Norte y EE.UU. Este bloque resulta fundamental para el traductor especializado, ya que el manejo de textología paralela se torna imprescindible cuando se está traduciendo cualquier tipo de documentación. Por esta razón, la autora incluye un abanico documental considerable: poderes para pleitos, poderes generales, poderes para administrar, poderes para vender, poderes para comprar, poderes para herencias, poderes para matrimonio civil, poderes para matrimonio religioso, poderes mercantiles y de sociedades, poderes conferidos por administrador, poderes acordados en junta, poderes “post mortem”, casos especiales de poderes, revocaciones de poderes, renuncias de poderes, poderes bilingües, declaraciones de voluntades anticipadas, declaraciones juradas y un largo etcétera.

El libro finaliza con tres anexos dedicados a ampliar los formularios, aportar legislación civil vigente y legislación notarial relevante. El anexo I enriquece el dosier de formularios expuesto anteriormente, aunque ahora se centra únicamente en los fragmentos más recurrentes de los mismos. El anexo II detalla la legislación pertinente en España, ofreciendo una selección de artículos del Código Civil español actualizado. Por su parte, el anexo III ofrece la normativa notarial pertinente: el “Reglamento Notarial”, muy útil para las autorizaciones de los notarios, la intervención de los testigos en el acto, el propio otorgamiento, la legitimación de firmas o las legalizaciones, entre otros aspectos.

El apartado de Bibliografía es muy completo y está totalmente actualizado, aspecto fundamental cuando se manejan textos jurídicos, teniendo en cuenta la variabilidad y versatilidad legislativa.

En esta obra, de carácter inédito, se observa un claro deseo de ofrecer a la comunidad científica una herramienta útil y práctica para el ejercicio profesional y docente. La traducción de documentación notarial en general y de los documentos de apoderamiento en particular constituye un encargo de traducción frecuente en la vida profesional. Sin embargo, la traducción de este tipo de documentos no está exenta de dificultades. Como bien indica su autora, los sistemas de derecho a estudio y cotejo son tangencialmente divergentes, con bases arraigadas de forma sustancialmente diversa: el ordenamiento jurídico español es prescriptivo y cuenta con una Constitución, mientras que el anglogermánico es descriptivo (y únicamente Estados Unidos dispone de documento constitucional per se). La base del derecho español es eminentemente romana (derecho continental), mientras que la base del derecho anglogermánico (derecho consuetudinario) es la denominada *Common Law*, basada en la costumbre y en el precedente.

A pesar de la dificultad que entraña el ámbito jurídico seleccionado por la autora, el libro se caracteriza por una redacción clara y precisa, pensada tanto para discentes como para docentes del ramo. Esta obra, hasta ahora única en el mundo, cubre un hueco esencial en el ámbito de la traducción especializada en general y la traducción jurídico-jurada en particular. A su vez, abarca cuestiones novedosas para el derecho de representación civil y penal comparado. En suma, el libro va destinado a varios tipos de lectores: traductores, juristas, lingüistas y terminólogos.

***

## The Journal is hosted by:

Department of International Business Communication (IBC)
Copenhagen Business School, Denmark

## Indexed in:

ERIH European Reference Index for the Humanities (European Science Foundation) – www.esf.org
Cambridge Scientific Abstracts (Linguistics and Language Behavior) – www.csa.com
MLA International Bibliography – www.mla.org
Norsk Samfunnsvitenskapelig Datatjeneste AS - http://dbh.nsd.uib.no
EBSCO Publishing Communications Sources – www.ebscohost.com (in indexing process)

## Contact:

The LSP journal
Copenhagen Business School
Dalgas Have 15
DK-2000 Frederiksberg, Denmark
Email: lsp@cbs.dk

## Electronic journal: **http://lsp.cbs.dk**

ISSN 1904-4135