A key perspective on specialised lexis: keywords in Telecommunication English for CLIL

Camino Rea Rizzo & María José Marín Pérez
Universidad Politécnica de Cartagena & Universidad de Murcia

7TH INTERNATIONAL CONFERENCE ON CORPUS LINGUISTICS (CILC2015)
Valladolid, 5-7 March 2015

A KEY PERSPECTIVE ON SPECIALISED LEXIS: KEYWORDS IN TELECOMMUNICATION ENGLISH FOR CLIL

Contents

- Introduction
  - CLIL, Tertiary Education and Corpus Linguistics
- Bilingual degree in Telecommunication Engineering and CLIL
  - Characteristics of the bilingual degree
  - CLIL hybrid approach: models C2 & C3
- Telecommunication Engineering Corpus (TEC)
  - Definition
  - Source of the samples
  - Topic representativeness
  - Structure

A KEY PERSPECTIVE ON SPECIALISED LEXIS: KEYWORDS IN TELECOMMUNICATION ENGLISH FOR CLIL

Contents

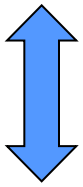
- Keywords
  - Keywords and terms
  - Keywords and their distribution
  - Keywords in an individual lesson: *Systems and circuits*
  - Training keywords
- Final remarks

Introduction

- **CLIL, Tertiary Education & Corpus Linguistics**
  - **C**ontent and **L**anguage **I**ntegrated **L**earning (CLIL)
    - CLIL aiming at acquiring a knowledge and command of at least two foreign languages (European Commission, 2003).
    - In 2009 more than 30 institutions in Spain offered bilingual degrees (Dafouz & Núñez).
    - In 2011 Business Administration at UMU and UPCT.
    - 2014/2015 Telecommunication Engineering at UPCT.
  - Language itself is also a learning goal: the vehicular language to convey content.
  - **T**elecommunication **E**nglish **C**orpus (TEC): academic and professional English

Introduction

- **CLIL, Tertiary Education & Corpus Linguistics**
  - CLIL **4Cs** conceptual framework: Content, Communication, Cognition and Culture (Coyle, Hood & Marsh, 2010)
    - **Language Triptych**: language ***of***, *for*, *through* learning.
      - Language **of** learning "explores what language learners will need to access new knowledge and understanding when dealing with the content."
      - The key vocabulary and phrases of the content language.



Language **of** learning = **keywords** in Wordsmith (Scott, 2008).

Bilingual degree in Telecom. Engineering & CLIL

Characteristics of the bilingual degree

- Telecommunication System Engineering **(TSE)** & Telematic Engineering **(TE)**
  - Goal: to improve students' competence in English while learning the specific content (easier access to labour market & further self-study).
  - 4Cs/Communication:
    - language is a conduit for communication and for leaning.
    - "language to use language and using language to learn" (Coyle et al., 2010).

% English	TSE	TE
1st year	50.5	50.5
2nd year	83	83
3rd year	75.4	70.5
4th year	75	86

Subjects in English	Basic, core, specific, compulsory and optional
100%	Lectures, bibliography, practicals, assignments, etc.
75%	Some lectures in Spanish
60%	Some lectures and practicals in Spanish
Technical English (one semester in 3rd year)	

Bilingual degree in Telecom. Engineering & CLIL

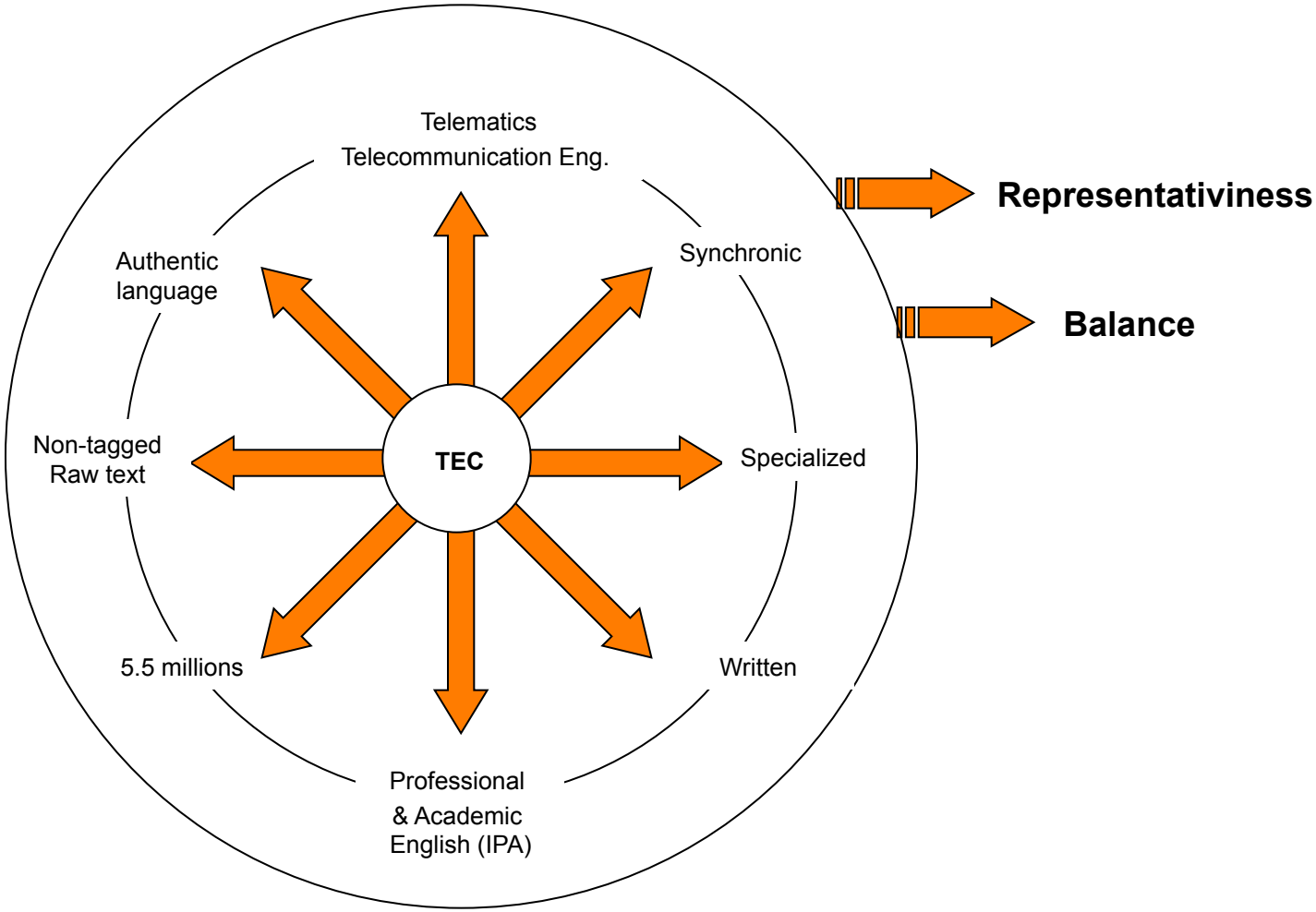
- **CLIL hybrid approach: models C2 & C3**
  - C2: Adjunct CLIL
    - "Language teaching runs parallel to content teaching with specific focus on developing the knowledge and skills to use the language so as to achieve higher-order thinking" (Coyle et al., 2010).
  - C3: Language embedded content courses
    - "Content programmes are designed from the outset with language development objectives. Teaching is carried out by content and language specialists "(ibid).
  - Trend in CLIL programs:
    - "include the teaching of the target language as a subject parallel to its being used as a vehicle for content-matter learning"(García, 2009).

Bilingual degree in Telecom. Engineering & CLIL

- **CLIL hybrid approach: models C2 & C3**
  - Language class
    - Language ***for*** learning: “the language needed by learners to operate in a language environment where the medium is not their first language” (Coyle et al., 2010) .
    - Language ***of*** learning, whose keywords and key phrases could be extracted from a specific corpus.
    - General characteristics of the sublanguage; pre-teaching or reinforcing the language of learning agreed with the content teacher.
  - Content class
    - Particular keywords of the lesson as a support of the language content.

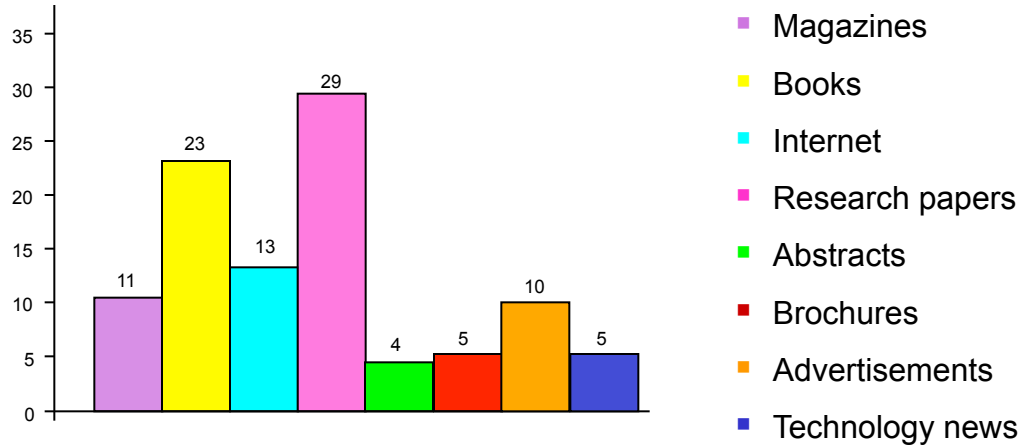
Telecommunication Engineering Corpus (TEC)

- **Definition**



TEC

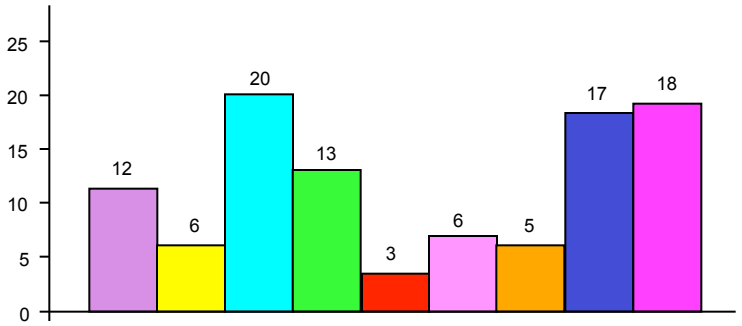
- **Source of the samples**



TEC

- **Topic representativeness**
  - **Thematic variety:** Curricula of Telecommunication Eng. & Telematics as a reference

→ 7 areas of knowledge + 2 majors



- Electronics
- Computing Architecture
- Telematic Engineering
- **Signal processing**
- Materials science
- Business management
- System engineering
- Communication networks
- Planning and management

TEC

Structure

Origin	Areas of knowledge	Subject areas	Sources
British	Electronics	Electronic components	Magazines
American	CAT	Electromagnetic fields	Books
Non native	Telematics	Business economy & manag.	Web
	Signal proc.	Analogue Electronics	Journals
	Materials	Digital electronics	Abstracts
	Business	Photonics	Brochures
	Systems	Computing fundamentals	Adverts
	*Com. Networks	Control engineering	News
	*Plan & Management	Instrumentation	
		TIC Materials	
		Projects	
		Telec. planning & management	
		Concurrent systems	
		Digital electronic systems	
		Distributed information systems	
		Systems & circuits	
		Systems & networks	
		Communication software	
		Telematics	
		Information processing	

Keywords

Keywords and terms

- Positive keywords given by WordSmith's tool (Scott, 2008):
  - Words whose frequency is unusually high in comparison to a general norm.
  - Words which are more probable to occur in telecommunications.
  - Words which usually provide a good account of the subject content.
- Keyword tool succeeds in identifying technical terms (Marín, 2014):
  - Even more accurate than other automatic term recognition methods (ATRM).
  - Keywords ranked 2nd out of 10 ATRMs, identifying 85% true terms out the top 400 candidate terms automatically extracted.

Keywords

Keywords and terms

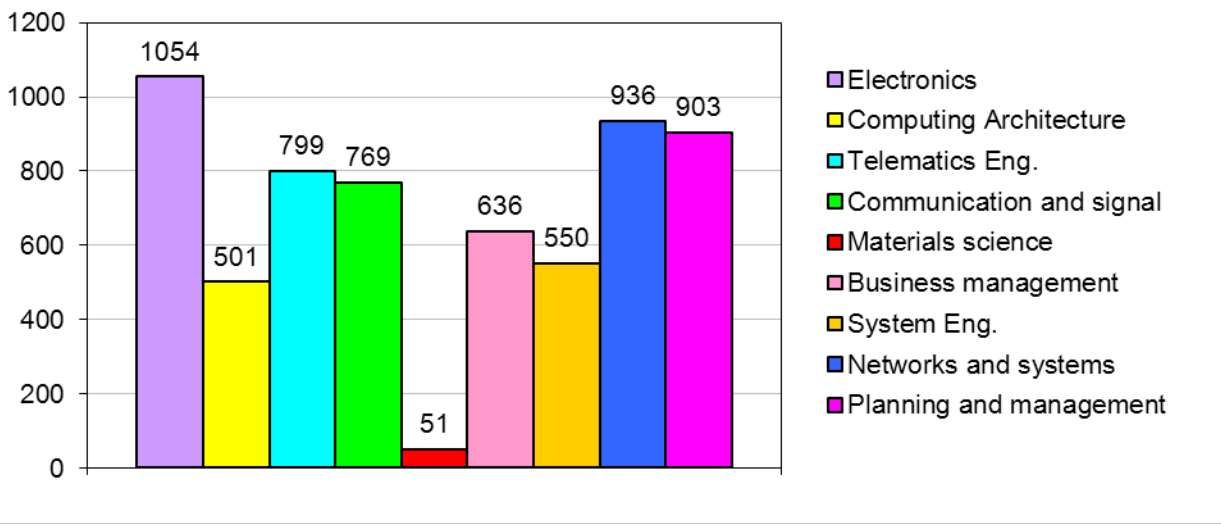
- Mastering terms is essential for successful communication:
  - A subject domain is not completely assimilated if the speaker is not familiar with its terminology.
  - A term entails a relative frequency > in technical than in general discourse BUT
  - It doesn't impose a high probability of occurrence in specialised texts.

Low frequency terms	High frequency terms
Peerware 21, encryptor 6, unicasting 6, bootable1, axially 1	Satellite 1401, VoIP 580, OSI 636, router 3910

- It is convenient to study first the most probable specialised lexical units that we may encounter independently of the degree of restriction to the discipline.

Keywords

- **Keywords and their distribution**
  - TEC compared to LACELL (20 millions / general English)
    - Reference keyword list: 5834 keywords (p value= 0)



Keywords

Keywords and their distribution

Keywords	Freq. TEC	Freq. LacellI	Key index	E	C.A	T	S	N	B	S	S.S	S.T
Network	16.649	1.686	41.784	-	-	3	-	-	-	-	-	802
Data	14.613	2.787	31.852	-	-	-	-	-	-	-	-	802
Systems	9.479	3.000	17.922	-	-	-	-	-	-	-	-	-
IP	5.239	20	17.377	-	-	3	-	-	-	-	-	802
Networks	5.832	463	16.182	-	-	3	-	-	-	-	801	802
System	12.624	8.707	15.204	-	-	-	-	-	-	-	-	-
Protocol	4.742	139	14.831	-	-	3	-	-	-	-	-	802
Design	7.701	3.313	14.725	1	2	-	4	-	-	-	-	-
Router	3.910	25	13.677	-	-	3	-	-	-	-	-	802
Wireless	4.083	171	12.237	-	-	-	4	-	-	-	801	802
Layer	4.425	569	12.117	-	-	-	-	-	-	-	-	802
Mobile	4.341	526	11.974	-	-	-	-	-	-	-	801	-
Input	4.347	709	11.914	1	2	-	4	-	-	-	-	-
Internet	4.504	910	11.589	-	-	3	-	-	-	-	-	802
Interface	3.526	207	11.454	-	-	3	-	-	-	-	-	802
Bandwidth	3.119	20	11.439	-	-	-	-	-	-	-	801	802
Packet	3.577	251	11.299	-	-	3	-	-	-	-	-	802
Circuit	3.932	525	10.804	1	2	-	4	-	-	-	-	-
Access	5.999	2.696	10.690	-	-	3	-	-	-	-	801	802
Output	4.139	771	10.604	1	2	-	4	-	-	-	-	-
Server	3.574	362	10.529	-	-	3	-	-	-	-	-	802
Digital	3.595	488	9.868	1	-	-	4	-	-	-	801	-
Software	4.575	1.412	9.860	-	2	3	4	-	-	-	-	-
Simulation	2.817	73	9.651	1	2	-	4	-	-	7	-	-
Devices	3.430	476	9.557	1	-	-	-	-	-	-	-	802
Voltage	2.945	220	9.551	1	-	-	-	-	-	-	-	-
Optical	2.822	164	9.485	1	-	-	-	5	-	-	801	-

- No word is key in all sections
- Top distribution value= 4 (*simulation, components, graph, quantum*)
- 3509 keywords distribution= 1
- 487 keywords distribution= 2
- 67 keywords distribution= 3

Keywords

- **Keywords and their distribution**
  - Restricted keywords

Areas	**Nº**	**Examples**
Electronics	570	Photoconductor, polarity, wavefront
Computing Architecture	223	Flops, microcontroller, caches
Telematics	558	Buffered, OGSI, applets, repository
Signal	357	Scintillation, bandpass, wavelets
Materials	150	Nanofibres, foams, tantalum
Business	200	Roamabout, teleworkers, globals
Systems	238	Debugger, pipeline, invariant, controllers
Sp. Signal	577	Layered, multiplexed, offline, WAP, GIS
Sp. Telematics	636	Modems, hackers, payload, unicast

Keywords

Keywords in an individual lesson

- 1st practical session of the subject System and Circuits, 1st year of TSE and TE

Practical 1: Basic instrumentation and passive components.

The first practical is a brief introduction of the main laboratory instruments with which the student will have to work when performing the generation and measurement of a given electrical quantity. They have to become familiar with the use of the laboratory equipment. A brief description of their main functions and different modes of operation will be provided in this practical. The student must himself practise with the equipment to acquire the necessary skills in handling.

Carefully read the contents of the practical before the laboratory session, both the descriptive part of each of the instruments and the exercises that are proposed to be carried out in the laboratory. This will lead to a better understanding of it and it will help to learn and achieve results in the practice session.

(1091 types)

Keywords

Keywords in an individual lesson

- 1st practical session of the subject System and Circuits, 1st year of TSE and TE
  - 63 keywords (out of 1901 types) are found in the text: *current, voltage, resistor, circuit, waveform, instrument, voltmeter, ohmmeter, sinusoidal, polymer, ripple,* etc.
  - *Current* top frequency in the practical: 29
    - Frequency in the whole corpus: 5064
    - Frequency in Communication & Signal Theory: 753
  - Patterns of *current*: 11 keywords from the practical co-occur with *current*.
  - Importance of mastering keywords:
    - Keywords occur frequently so there's ample opportunity to meet and use them.
    - Recurrent exposure to keywords contributes to consolidate knowledge.

Keywords

- **Exercises to train keywords**
  - Aspects involved in knowing a word: form, meaning and use (Nation, 2001)
    - Data-driven learning experiments presented by Boulton (20100) as a reference for vocabulary-focused activities.
    - Marín (2014) focuses on the morphological, syntactic, semantic and discursive levels and designs specific exercises:
      - Reflect on the process of word formation
      - Identify lexical and grammatical patterns associated with a particular term
      - Develop a written project

Final remarks

- **Fruitful relationship between Corpus Linguistics & CLIL**
  - Language ***of*** learning = keywords
    - Related to the specific domain/terms;
    - Recurrent: worth studying;
    - Frequent: consolidate knowledge;
    - Useful for understanding the content subject.
- **BUT**
  - Language data need processing and human supervision
    - CLIL approach demands a strong collaboration between language teachers and content-subject teachers.

Bilingual degrees entails a different approach and teaching methodology where there is little doubt for the adequacy and profitability of a specific corpus.

A key perspective on specialised lexis: keywords in Telecommunication English for CLIL

Camino Rea Rizzo & María José Marín Pérez
Universidad Politécnica de Cartagena & Universidad de Murcia

7TH INTERNATIONAL CONFERENCE ON CORPUS LINGUISTICS (CILC2015)
Valladolid, 5-7 March 2015