

UNIVERSIDAD DE MURCIA
FACULTAD DE LETRAS



Departamento de Filología Inglesa

TESIS DOCTORAL

**EL INGLÉS DE LAS TELECOMUNICACIONES:
ESTUDIO LÉXICO BASADO EN UN CORPUS ESPECÍFICO**

Presentada por Camino Rea Rizzo

Dirigida por el Dr. Aquilino Sánchez Pérez,
Catedrático de la Universidad de Murcia

2008

A mi abuelo Carmelo

AGRADECIMIENTOS

Mi más sincero agradecimiento al profesor Sánchez por su inestimable ayuda en la dirección de esta tesis, y por la confianza y el apoyo que siempre he recibido de él.

Me gustaría extender mi agradecimiento a todos aquellos que, directa o indirectamente, han aportado su granito de arena en los diferentes aspectos y etapas que supone la realización de un trabajo de tal envergadura. Entre estas personas quisiera mencionar a Mari Carmen, por su apoyo estadístico y personal desde el primer día; a los profesores de la Escuela de Ingeniería de Telecomunicación de la Universidad Politécnica de Cartagena que me han facilitado materiales y referencias de gran utilidad; a mi compañera de despacho, por su colaboración especialmente enriquecedora y su apoyo desinteresado, no sólo en el ámbito académico.

Sin duda alguna, no podría haberme hallado en esta situación sin el esfuerzo realizado por mis padres, a quienes agradezco todas las oportunidades que me han brindado en la vida.

Finalmente, quiero agradecer la constante presencia y el apoyo incondicional de Javi, que apostó por compartir su vida conmigo. Él ha ido descifrando en cada momento la etapa en la que me encontraba, a la vez que señalaba, con un gesto de cariño, lo que me quedaba para terminar.

ÍNDICE DE CONTENIDO

CAPÍTULO 1. INTRODUCCIÓN	1
1.1 CONTEXTUALIZACIÓN Y OBJETIVOS GENERALES.	1
1.2 OBJETIVOS ESPECÍFICOS Y PREGUNTAS DE LA INVESTIGACIÓN.	3
1.3 RESUMEN DE CONTENIDOS POR CAPÍTULOS.	4
CAPÍTULO 2. LOS LENGUAJES DE ESPECIALIDAD	7
2.1 COMUNICACIÓN GENERAL Y COMUNICACIÓN ESPECIALIZADA.	7
2.2 DEFINICIÓN DE LENGUAJES DE ESPECIALIDAD.	10
2.3 EL LÉXICO DE LOS LENGUAJES DE ESPECIALIDAD.	12
CAPÍTULO 3. EL INGLÉS CON FINES ESPECÍFICOS: EL INGLÉS PROFESIONAL Y ACADÉMICO	15
3.1 INICIOS Y DEFINICIÓN.	15
3.2 EVOLUCIÓN DEL IPA.	17
3.2.1 <i>Análisis del registro.</i>	19
3.2.2 <i>Análisis retórico o análisis del discurso.</i>	19
3.2.3 <i>Análisis de la situación término (Target situation analysis).</i>	20
3.2.4 <i>Destrezas y estrategias.</i>	20
3.2.5 <i>Enfoque centrado en el aprendizaje (Learning-centred approach).</i>	21
3.2.6 <i>Análisis del Género.</i>	21
3.2.7 <i>Tendencias actuales y líneas de investigación.</i>	23
3.3 ANÁLISIS DE NECESIDADES.	25
3.4 CONTEXTO METODOLÓGICO. EL ANÁLISIS DEL TEXTO SEGÚN LA TRADICIÓN BRITÁNICA:	
FIRTH, HALLIDAY Y SINCLAIR.	29
<i>Principio 1: La Lingüística es esencialmente una ciencia social y aplicada.</i>	30
<i>Principio 2: El lenguaje se debería estudiar en muestras auténticas, acreditadas y reales, y no por medio de frases aisladas, inventadas e intuitivas.</i>	31
<i>Principio 3: La unidad de estudio debe ser el texto completo.</i>	31
<i>Principio 4: Los textos y tipos de texto se deben estudiar de forma comparada por medio de corpus textuales.</i>	32
<i>Principio 5: La Lingüística se ocupa del estudio del significado: forma y significado son inseparables.</i>	33
<i>Principio 6: No existen fronteras entre léxico y gramática: léxico y gramática son interdependientes.</i>	33
<i>Principio 7: Gran parte del lenguaje es rutina.</i>	34
<i>Principio 8: La lengua en uso transmite la cultura.</i>	34
<i>Principio 9: Los dualismos de Saussure están mal planteados.</i>	34
CAPÍTULO 4. LA LINGÜÍSTICA DEL CORPUS	37
4.1 INICIOS Y EVOLUCIÓN DE LA LINGÜÍSTICA DEL CORPUS.	37
4.2 DEFINICIÓN Y TIPOLOGÍA DE CORPUS LINGÜÍSTICO.	41
4.2.1 <i>Definición de corpus lingüístico.</i>	41
4.2.2 <i>Tipos de corpus.</i>	42
4.3 CORPUS PRINCIPALES.	43
4.3.1 <i>Corpus de primera generación.</i>	44
<i>The Brown Corpus: The Standard Sample of Present-Day American English.</i>	44
<i>The London-Lund Corpus of Spoken English (LLC).</i>	44
<i>The Lancaster-Oslo-Bergen Corpus of British English (LOB).</i>	45

ÍNDICE

4.3.2	<i>Corpus de segunda generación.</i>	45
	<i>The Longman-Lancaster English Language Corpus.</i>	45
	<i>Proyecto Cobuild.</i>	46
4.3.3	<i>Corpus de 3ª generación.</i>	46
	<i>The Bank of English Corpus.</i>	46
	<i>The British National Corpus (BNC).</i>	47
	<i>The International Corpus of English (ICE).</i>	47
	<i>The American National Corpus (ANC).</i>	47
4.4	CORPUS ESPECIALIZADOS.	48
4.5	CONSIDERACIONES RELEVANTES PARA LA CREACIÓN DE UN CORPUS.	51
4.5.1	<i>Representatividad.</i>	52
4.5.2	<i>Tamaño.</i>	53
4.5.3	<i>Variedad.</i>	54
4.5.4	<i>Cronología.</i>	54
4.5.5	<i>Tipos de texto o géneros.</i>	54
4.5.6	<i>Anotación del corpus.</i>	55
CAPÍTULO 5. DISEÑO Y DISTRIBUCIÓN DEL CORPUS DE TELECOMUNICACIONES		57
5.1	DISEÑO DEL CORPUS.	57
5.1.1	<i>Variedad temática.</i>	57
5.1.2	<i>Cronología.</i>	58
	Fase I.	58
	Fase II.	58
5.1.3	<i>Ámbito Geográfico.</i>	58
5.1.4	<i>Modalidad.</i>	59
5.1.5	<i>Tamaño.</i>	59
5.2	DISTRIBUCIÓN DEL CORPUS.	60
CAPÍTULO 6. APLICACIONES INFORMÁTICAS PARA EL PROCESAMIENTO DEL CORPUS		67
6.1	SOFTWARE PARA EL PROCESAMIENTO DEL CORPUS.	67
6.2	INFORMACIÓN ESTADÍSTICA BÁSICA SOBRE LA COMPOSICIÓN DEL CORPUS.	68
6.2.1	<i>Listados de frecuencias: comparación general de los dos corpus.</i>	71
CAPÍTULO 7. ESTUDIO DEL VOCABULARIO		77
7.1	JUSTIFICACIÓN DEL ESTUDIO DEL VOCABULARIO TÉCNICO DESDE UNA PERSPECTIVA DIDÁCTICA.	77
7.2	CLASIFICACIÓN DE LOS DIFERENTES TIPOS DE VOCABULARIO EN UN TEXTO ESPECIALIZADO POR MEDIO DE CRITERIOS CUALITATIVOS.	78
7.3	CLASIFICACIÓN DE LOS DIFERENTES TIPOS DE VOCABULARIO EN UN TEXTO ESPECIALIZADO POR MEDIO DE CRITERIOS CUANTITATIVOS Y CORPUS LINGÜÍSTICOS.	81
7.4	COMBINACIÓN DE CRITERIOS.	87
7.4.1	<i>Vocabulario general.</i>	87
7.4.2	<i>Vocabulario académico.</i>	88
7.4.3	<i>Vocabulario técnico.</i>	88
7.4.4	<i>Vocabulario semitécnico.</i>	89
7.5	MÉTODOS PARA SELECCIONAR LOS TIPOS VOCABULARIO.	89
7.5.1	<i>Método de listas de corte.</i>	89
7.5.2	<i>Método de comparación de corpus.</i>	90
CAPÍTULO 8. ANÁLISIS DEL CORPUS Y RESULTADOS		91
8.1	CLASIFICACIÓN DEL VOCABULARIO: APLICACIÓN DEL MÉTODO DE LISTAS DE CORTE.	91
8.1.1	<i>Listas de corte y filtrado del corpus.</i>	91

ÍNDICE

8.1.1.1 Elección de las listas de corte.	92
8.1.1.2 Filtrado del corpus completo.	93
8.1.2 <i>Distribución de las palabras por áreas de conocimiento.</i>	94
8.1.2.1 Filtrado del corpus por áreas de conocimiento.	94
8.1.2.2 Contraste de los listados.	94
8.1.2.3 Palabras únicas de cada área.	101
8.2 CLASIFICACIÓN DEL VOCABULARIO MEDIANTE COMPARACIÓN DE CORPUS.	105
8.2.1 <i>Cálculo de la frecuencia relativa y el ratio.</i>	106
8.2.2 <i>Agrupación de las palabras con características similares.</i>	107
8.2.2.1 Grupo 1: Palabras que sólo aparecen en el específico y ocurren sólo una vez.	109
8.2.2.2 Grupo 2: Palabras que sólo aparecen en el específico y con frecuencia ≥ 2 .	109
8.2.2.3 Grupo 3: Palabras que aparecen en los dos corpus, su frecuencia en el corpus técnico es ≥ 12 y el ratio es ≥ 50 .	110
8.2.2.4 Grupo 4, grupo 5, grupo 6 y grupo 7.	110
8.2.2.5 Palabras con frecuencia 1 en el corpus general.	112
8.2.2.6 Terminología.	112
8.2.2.7 Terminología generalizada	113
8.2.3 <i>Conclusión.</i>	113
8.3 KEYWORDS O PALABRAS CLAVE DEL CORPUS.	114
8.3.1 <i>Listados de Keywords.</i>	115
8.3.2 <i>Distribución de las keywords por áreas de conocimiento.</i>	118
8.3.3 <i>Key-keywords.</i>	122
CAPÍTULO 9. FAMILIAS DE PALABRAS TÉCNICAS	125
9.1 INTRODUCCIÓN.	125
9.2 PROCESO DE SELECCIÓN DE LAS FAMILIAS DE PALABRAS TÉCNICAS.	127
9.2.1 <i>Criterios preliminares.</i>	127
9.2.2 <i>Familia de palabras: Agrupación de miembros.</i>	129
9.2.3 <i>Marcado de palabras académicas y generales más frecuentes.</i>	131
9.3 RESULTADO DE LA CLASIFICACIÓN DE FAMILIAS EN EL CORPUS DE TELECOMUNICACIONES.	136
9.3.1 <i>Familias de palabras técnicas.</i>	136
Familias calificadas como NO específicas.	136
Palabras etiquetadas como "WEST TÉCNICA".	137
Formas abreviadas.	138
Familias en las que todos los miembros son valorados como específicos.	139
Familias en las que más de la mitad de los miembros son valorados como específicos.	140
Resto de familias.	141
9.3.2 <i>Familias de palabras generales.</i>	142
9.3.3 <i>Familias de palabras académicas.</i>	144
9.4 CONCLUSIÓN Y CRITERIOS FINALES.	145
CAPÍTULO 10. PALABRAS CON PALABRAS: ANÁLISIS DETALLADO	149
10.1 INTRODUCCIÓN.	149
10.2 DESARROLLO DEL ANÁLISIS DETALLADO: BANDWIDTH.	150
10.2.1 <i>Frecuencia.</i>	150
10.2.2 <i>Distribución.</i>	151
10.2.3 <i>Colocación y colocados significativos.</i>	152
10.2.4 <i>Agrupaciones léxicas o clusters.</i>	161
10.3 ANÁLISIS DETALLADO DE FORMAS TÉCNICAS.	166
10.4 ANÁLISIS DETALLADO DE FORMAS ACADÉMICAS.	225
10.5 ANÁLISIS DETALLADO DE FORMAS GENERALES.	252
10.6 CONCLUSIÓN.	313

ÍNDICE

CAPÍTULO 11. EL VOCABULARIO ESPECIALIZADO EN LOS TEXTOS ESPECÍFICOS	323
11.1 VALIDACIÓN DEL REPERTORIO DE VOCABULARIO ESPECIALIZADO.	323
11.2 EL VOCABULARIO ESPECIALIZADO EN LOS LIBROS DE TEXTO DE INGLÉS DE LA ESPECIALIDAD.	326
11.2.1 <i>Contraste de listados de palabras clave.</i>	327
11.3 EL REPERTORIO DE VOCABULARIO ESPECIALIZADO EN LOS LIBROS DE TEXTO.	331
CAPÍTULO 12. CONCLUSIONES, APLICACIONES Y FUTUROS TRABAJOS	335
BIBLIOGRAFÍA	341
APÉNDICES	CD

ÍNDICE DE FIGURAS

<i>Figura 2.1 Modelo de comunicación según Jakobson (1960).</i>	8
<i>Figura 2.2 Relación entre lengua global, lengua común y lenguajes de especialidad.</i>	10
<i>Figura 5.1 Volumen del corpus distribuido por variedades geográficas.</i>	61
<i>Figura 5.2 Volumen del corpus distribuido por áreas temáticas.</i>	63
<i>Figura 5.3 Volumen del corpus distribuido por tipos de texto.</i>	65
<i>Figura 6.1 Relación entre el número de palabras, formas y lemas.</i>	71
<i>Figura 8.1 Volumen de palabras obtenido en cada grupo.</i>	114
<i>Figura 9.1 Características de las familias candidatas a términos.</i>	141
<i>Figura 9.2 Comportamiento de las familias generales.</i>	143
<i>Figura 11.1 Relación formas especializadas – formas nocionales – tokens.</i>	334
<i>Figura 11.2 Fracción del repertorio de vocabulario especializado en los libros.</i>	334

ÍNDICE DE CUADROS

<i>Cuadro 4.1 Otros corpus en lengua inglesa.</i>	48
<i>Cuadro 5.1 Volumen del corpus distribuido por variedades geográficas</i>	60
<i>Cuadro 5.2 Código completo de ATC de procedencia británica</i>	66
<i>Cuadro 6.1 Información estadística del Corpus de Telecomunicaciones.</i>	69
<i>Cuadro 6.2 Información estadística del Corpus Lacell.</i>	70
<i>Cuadro 6.3 Proporción de corpus que cubren las palabras más frecuentes.</i>	73
<i>Cuadro 6.4 Densidad léxica</i>	74
<i>Cuadro 8.1 Clasificación cualitativa en función de la frecuencia y distribución.</i>	105
<i>Cuadro 8.2 Clasificación cualitativa en función de la frecuencia y distribución.</i>	105
<i>Cuadro 8.3 Resumen del volumen de palabras obtenido en cada grupo.</i>	113
<i>Cuadro 8.4 Volumen de Keywords $p = 0$.</i>	119
<i>Cuadro 8.5 Proporción de keywords.</i>	121
<i>Cuadro 8.6 Keywords únicas de cada área.</i>	122
<i>Cuadro 9.1 Tamaño de vocabulario y porcentaje de cobertura en novelas para adolescentes.</i>	125
<i>Cuadro 9.2 Cobertura del vocabulario.</i>	127
<i>Cuadro 9.3 Etiquetadas general/académica, valoradas como específicas.</i>	134
<i>Cuadro 9.4 Incidencia de familias.</i>	135
<i>Cuadro 9.5 Siglas</i>	139
<i>Cuadro 9.6 Proporción de formas y su cobertura del corpus.</i>	147
<i>Cuadro 10.1 Concordancias de memory bandwidth.</i>	160
<i>Cuadro 10.2 Clusters de diferentes unidades de bandwidth.</i>	161
<i>Cuadro 11.1 Validez de la lista en textos del corpus.</i>	324
<i>Cuadro 11.2 Validez de la lista en textos ajenos al corpus.</i>	325
<i>Cuadro 11.3 Diversidad léxica.</i>	326
<i>Cuadro 11.4 Información estadística básica de los libros de texto.</i>	327

ÍNDICE DE TABLAS

<i>Tabla 3.1 Etapas en la evolución del IPA.</i>	18
<i>Tabla 6.1 Listado de las 200 palabras más frecuentes en el corpus técnico y en el general.</i>	73
<i>Tabla 6.2 Las 50 palabras de contenido más frecuentes.</i>	76
<i>Tabla 8.1 Distribución de las palabras por áreas de conocimiento.</i>	97
<i>Tabla 8.2 Palabras presentes en todas las áreas.</i>	101
<i>Tabla 8.3 Volumen de palabras únicas de cada área.</i>	102
<i>Tabla 8.4 Cobertura de las palabras únicas.</i>	102
<i>Tabla 8.5a Únicas del área.</i>	103
<i>Tabla 8.5b Únicas del área.</i>	103
<i>Tabla 8.6 Keywords positivas de Teleco</i>	117
<i>Tabla 8.7 Distribución de keywords por áreas de conocimiento.</i>	121
<i>Tabla 8.8 Key-keywords.</i>	124
<i>Tabla 9.1 Listado base para familias técnicas.</i>	128
<i>Tabla 9.2 Familias de palabras.</i>	131
<i>Tabla 9.3 Marcado de las familias.</i>	134
<i>Tabla 9.4 Familias etiquetadas como WEST TÉCNICA.</i>	138
<i>Tabla 10.1 Familia de band.</i>	151
<i>Tabla 10.2 Distribución de bandwidth.</i>	152
<i>Tabla 10.3 Colocados de bandwidth.</i>	154
<i>Tabla 10.4 Colocados significativos de bandwidth.</i>	157
<i>Tabla 10.5 Colocados contiguos de las combinaciones especializadas.</i>	165
<i>Tabla 10.6 El vocabulario especializado de la ingeniería de telecomunicación. (Las 1.000 formas más significativas).</i>	322
<i>Tabla 11.1 Keywords de los libros de texto por separado.</i>	328
<i>Tabla 11.2 Keywords</i>	331
<i>Tabla 11.3 Vocabulario especializado en los libros de texto.</i>	333

CAPÍTULO 1

INTRODUCCIÓN

1.1 Contextualización y objetivos generales.

El presente trabajo está enmarcado en la línea de investigación, propuesta por el Departamento de Filología Inglesa de la Universidad de Murcia, Estudio diacrónico y sincrónico del inglés mediante corpus lingüísticos. Más concretamente, surge con la pretensión de aportar posibles soluciones a los problemas que plantea la enseñanza y aprendizaje de la lengua inglesa en la Universidad Politécnica de Cartagena (UPCT). La identificación del problema nace de la experiencia docente en inglés específico para estudiantes de Ingeniería de Telecomunicación en la UPCT.

En 1999 la UPCT incorpora a sus centros de enseñanzas tecnológicas la Escuela Técnica Superior de Ingeniería de Telecomunicación, implantando las titulaciones de Ingeniero de Telecomunicación e Ingeniero Técnico de Telecomunicación, especialidad en Telemática. Como complemento de los planes de estudios, se crea una asignatura de inglés con fines específicos y no de inglés general. Durante la elaboración del programa para el primer curso (1999/2000), se advirtió la carencia de un material específico, adecuado y experimentado, que cubriera las necesidades lingüísticas de los estudiantes. Los libros de texto existentes, o estaban anticuados, o no guardaban relación con los contenidos que se pretendía enseñar, de modo que no eran aptos para tomarlos como material de referencia. Por tal razón, se decidió diseñar un programa propio, confiando en un principio en la intuición de los docentes para elegir temas, léxico, competencias y destrezas en las que centrar la enseñanza, así como en las experiencias docentes llevadas a cabo en contextos similares en otros centros.

Siendo conscientes del problema, se decidió emprender un estudio sustentado empíricamente: investigar qué deben saber los estudiantes de Ingeniería de Telecomunicaciones para poder desenvolverse en inglés en su futuro profesional y durante su formación académica, es decir, para poder adquirir, desarrollar y transmitir conocimientos especializados en la lengua inglesa de forma eficiente. Despejar dicha incógnita constituye el objetivo principal de la presente investigación.

El análisis de la lengua meta y de las necesidades del grupo particular de aprendices conduce a identificar la información requerida, y de este modo, a elaborar un programa

docente adecuado para la enseñanza y aprendizaje del inglés técnico de telecomunicaciones. Con ello, se persigue la mejora en la calidad de la enseñanza, ya que, desde el inicio de la asignatura, probablemente no se haya cubierto la totalidad de las áreas temáticas o de conocimiento de estos estudios ni sus situaciones profesionales. Por tanto, se pretende identificar qué aspectos del lenguaje como estructura y como herramienta de comunicación son relevantes, más importantes o útiles en tales contextos. Con este fin, surge la idea fundamental de la elaboración de un corpus lingüístico de donde extraer la información requerida y poder servir de base de investigación lingüística.

El marco teórico que encuadra esta investigación enlaza varias disciplinas, cuyos fundamentos constituyen una sólida base de referencia para examinar el objeto de estudio, y de donde se extraen las directrices para aplicar una metodología de análisis.

La pretensión de cubrir las necesidades comunicativas del alumnado mediante la creación del corpus, suscita el planteamiento de una serie de cuestiones que salen a la luz en un proceso de análisis de necesidades, especialmente las relacionadas con el entorno lingüístico académico y profesional, y que se abordarán desde la óptica de la Enseñanza de Lenguas con Fines Específicos. En dicho análisis se seguirán las recomendaciones de Dudley-Evans y St. John (1998), con la intención de obtener una serie de respuestas que orienten la recogida de muestras apropiadas para los fines de la investigación.

El diseño del corpus y la recopilación de datos se realiza siguiendo los criterios que dicta la Lingüística del Corpus, tomando muestras reales del uso del lenguaje como fuente de evidencia para la argumentación lingüística. Se incorporan metodologías que procesan el lenguaje por medios informáticos, cuantifican sus unidades y requieren la combinación de técnicas computacionales con interpretaciones cualitativas.

Asimismo, se adopta la perspectiva de la lexicología para describir el lenguaje especializado de las telecomunicaciones, ya que la principal diferencia entre la lengua general y la específica radica en el vocabulario que los hablantes utilizan para comunicarse. Además, la finalidad principal de la lexicología es describir las palabras de la lengua y su funcionamiento en el léxico del hablante.

Sin embargo, las unidades léxicas en estudio forman parte de un lenguaje de especialidad. El hecho de que el corpus compilado a partir de los planteamientos formulados refleje el discurso de una comunidad específica de hablantes, conlleva afrontar los términos específicos utilizados para transmitir conocimiento. En la comunicación especializada, los interlocutores utilizan la sintaxis, la morfología y una parte del léxico de la lengua general, más el subcódigo propio de la especialidad, que incluye la terminología específica de la materia en cuestión. No obstante, el objetivo de

esta investigación no es la búsqueda de términos para fijar su definición, aunque el estudio también pueda calificarse como terminológico, en el sentido de que se trata una variedad del lenguaje caracterizada por una especialización temática.

El conjunto de rasgos que definen la situación y el planteamiento del problema determinan que el tipo de investigación propuesta debe ser descriptivo. Mediante el análisis del objeto de estudio se pretende caracterizar la realidad lingüística de una comunidad discursiva, desde un enfoque léxico. Igualmente, los resultados obtenidos presentan una aplicación eminentemente práctica, puesto que la investigación está destinada a cubrir una carencia en la enseñanza del lenguaje especializado.

Por último, es pertinente añadir que el carácter multidisciplinar de esta investigación pone en contacto diferentes áreas, como la Lingüística del Corpus, la Enseñanza de Lenguas con Fines Específicos, la Lexicología, la Terminología y las Telecomunicaciones.

1.2 Objetivos específicos y preguntas de la investigación.

En el apartado anterior se ha expuesto de forma genérica el objetivo de la presente investigación, cuya finalidad consiste en identificar el contenido léxico requerido para poder adquirir, desarrollar y transmitir conocimientos especializados de la ingeniería de telecomunicación en lengua inglesa, tanto en un ámbito profesional como de formación académica. De forma más precisa, el objetivo principal se centra en concretar el repertorio léxico necesario para conseguir una comunicación efectiva en un contexto académico y profesional.

Este objetivo general permite formular la pregunta de investigación principal: ¿Qué léxico se requiere para poder adquirir, desarrollar y transmitir eficazmente en inglés conocimientos específicos de la ingeniería de telecomunicaciones?

Para dar respuesta a la pregunta de investigación principal y posibilitar la consecución del objetivo general, éste se ha desglosado en una serie de acciones que se corresponden con los siguientes objetivos específicos:

1. La recopilación de un corpus lingüístico del inglés propio de la ingeniería de telecomunicaciones.
2. La descripción del perfil léxico del inglés propio de las telecomunicaciones mediante:
 - la exploración de los diferentes tipos de vocabulario: general, técnico y semitécnico;

- la identificación de las palabras estadísticamente significativas, denominadas *Keywords*.
3. El análisis del comportamiento del léxico identificado.
 4. La identificación de las 1000 familias de palabras técnico-profesionales que contribuyen a conseguir una comprensión del 95% de los textos especializados.

Al igual que el objetivo principal, los objetivos específicos están asociados a una serie de preguntas que encauzan el desarrollo de la investigación. El conjunto de objetivos específicos y preguntas secundarias, que se enuncian a continuación, pretenden facilitar la consecución del objetivo general planteado.

¿El léxico de las telecomunicaciones es diferente del léxico de la lengua general? ¿Cuál es el perfil léxico del inglés de las telecomunicaciones? ¿Qué tipos de palabras se encuentran en el registro específico? ¿Qué palabras son representativas de este ámbito? ¿Cuál es el uso del léxico de las telecomunicaciones? ¿Qué asociaciones o agrupaciones de palabras tienen lugar normalmente? ¿Qué palabras suelen co-ocurrir? ¿Cuál es el vocabulario requerido para manejar los textos académicos y profesionales del inglés de la ingeniería de telecomunicaciones?

Después de caracterizar el perfil léxico y determinar el contenido léxico requerido, se procede a comprobar si los libros de texto de inglés para la ingeniería de telecomunicaciones incluyen el contenido léxico detectado y reflejan adecuadamente tal perfil léxico: ¿Existen diferencias significativas entre el vocabulario que se encuentra en los libros de texto de inglés y el utilizado realmente en ingeniería de telecomunicaciones?

1.3 Resumen de contenidos por capítulos.

Una vez que se ha expuesto la motivación del presente estudio, ubicado principalmente en el marco de la enseñanza de Lenguas con Fines Específicos y en la Lingüística del Corpus, y se han detallado los objetivos perseguidos, centrando la investigación en el nivel léxico del lenguaje de la ingeniería de telecomunicaciones, se procede a presentar la organización del trabajo y a resumir el contenido del resto de los capítulos que lo conforman.

La estructura que configura este estudio tiende a mostrar una perspectiva general y por separado de las disciplinas que lo fundamentan, para ir evolucionando hacia una visión más particular que especifica el objeto de estudio y que establece lazos de conexión entre las diferentes disciplinas. De este modo, se desarrollarán los fundamentos teóricos sobre los que se asienta la metodología de análisis y se seleccionarán los recursos que sirvan mejor a los propósitos de la investigación,

estableciendo así una base teórico-práctica enriquecida por una organización multidisciplinar. Después, se aplicará la metodología señalada y se procederá al análisis, descripción e interpretación de los datos obtenidos, con el fin de dar respuesta al conjunto de preguntas planteadas que rigen el curso de la investigación. Por último, se someterán a prueba los hallazgos conseguidos.

Tras el presente capítulo introductorio, el contenido del resto de los capítulos queda de la siguiente forma:

El capítulo 2 delimita el cuerpo de análisis del estudio: el lenguaje especializado de la ingeniería de telecomunicación. Partiendo de la concepción del lenguaje como sistema y como instrumento de comunicación, desde la lingüística teórica hasta la aplicada, se accederá al plano léxico y se describirá la relación existente entre la lengua general y los lenguajes de especialidad, y entre los diferentes lenguajes de especialidad.

El capítulo 3 define el inglés específico como área de conocimiento dentro de la enseñanza de lenguas extranjeras y realiza una revisión de las diferentes etapas que se han sucedido a lo largo de su historia, hasta llegar a las tendencias y líneas de investigación actuales. Asimismo, se hace el análisis de necesidades preciso, previo a la elaboración de un curso de inglés profesional y académico, considerando la situación en la que se utilizará el lenguaje especializado, la situación del aprendizaje y la situación de los alumnos. Teniendo en cuenta las necesidades detectadas y la carencia de material didáctico apropiado, se plantea la idoneidad de un corpus lingüístico para solucionar los problemas surgidos. El capítulo cierra con el contexto metodológico, concepción del lenguaje y principios teóricos que fundamentan la corriente lingüística donde se ubican los corpus lingüísticos.

El capítulo 4 enlaza directamente con el anterior, donde se exponen los inicios de la lingüística del corpus en relación con los estudios léxicos y cómo aquella evoluciona paralelamente al desarrollo de los ordenadores, dando lugar a diferentes tipos de corpus. Se pone de relieve la carencia de un corpus especializado y útil para el objetivo de la investigación, concluyendo con la necesidad de recopilar uno propio. Después, se detallan los criterios fundamentales que se deben seguir en el diseño y recopilación de un corpus.

El capítulo 5 describe las tareas de diseño, recopilación y distribución del corpus técnico de telecomunicaciones, de acuerdo con las directrices establecidas.

Seguidamente, el capítulo 6 introduce las aplicaciones informáticas y conceptos básicos para el procesamiento del corpus, aportando los primeros resultados estadísticos sobre el corpus técnico y el corpus general de referencia.

El capítulo 7 comienza con una justificación del estudio del vocabulario técnico desde una perspectiva didáctica. A continuación, presenta una clasificación de los diferentes tipos de vocabulario y las características que los definen, según criterios cualitativos y cuantitativos. Finalmente, se muestran dos métodos para la identificación del vocabulario basados en corpus lingüísticos, que facilitan además, su combinación con criterios de selección cuantitativos y cualitativos.

Tras el desarrollo y exposición de los fundamentos teóricos de la investigación y el planteamiento de los métodos de reconocimiento del vocabulario, el capítulo 8 expone el proceso de aplicación de la metodología que permite analizar el corpus y obtener los datos necesarios para caracterizar el perfil léxico, en función de los parámetros de restricción y representatividad de la unidad léxica en el dominio especializado.

El capítulo 9 trata de la relación existente entre los repertorios léxicos definidos (académico y general más frecuente) y la proporción de palabras que cubren en un texto. Ante la carencia de un repertorio establecido para el vocabulario especializado de la ingeniería de telecomunicaciones, se proponen, se prueban y se ajustan los criterios que permiten seleccionar dicho repertorio, haciendo uso de los resultados obtenidos en los análisis precedentes.

El capítulo 10 desarrolla un análisis detallado de una muestra de palabras seleccionadas entre las más significativas del vocabulario especializado. Los diferentes factores que intervienen en el análisis exploran el contexto inmediato de cada palabra y describen las relaciones semánticas que se establecen en el nivel sintagmático, con el fin de mostrar el comportamiento léxico en el discurso especializado.

La inclusión de un apartado dedicado a las conclusiones esgrimidas en los capítulos 8, 9, y 10, responde al deseo de presentar una progresión lógica del estudio, valorando la información obtenida: se recopilan los hallazgos conseguidos y se indican las carencias detectadas, lo que constituye la base del siguiente estadio.

El capítulo 11 contiene los resultados obtenidos tras comprobar la validez del repertorio especializado, e incluye el cotejo del contenido léxico de los libros de texto utilizados en la enseñanza del inglés técnico con dicho repertorio.

Finalmente, el capítulo 12 recoge las principales conclusiones, implicaciones y aplicaciones de este trabajo, al tiempo que ofrece nuevas vías de investigación.

La tesis doctoral se cierra con la sección dedicada a la bibliografía que ha hecho posible su elaboración. Por último, y por cuestiones de espacio, se ha decidido incluir en un CD los apéndices que completan y facilitan la lectura y comprensión del presente trabajo.

CAPÍTULO 2

LOS LENGUAJES DE ESPECIALIDAD

En esta sección se delimita el cuerpo de análisis del estudio: el lenguaje especializado de la ingeniería de telecomunicación. La caracterización de los lenguajes de especialidad se realizará desde diferentes ángulos, partiendo de la concepción del lenguaje y la comunicación lingüística bajo la óptica de la lingüística teórica y la lingüística aplicada, hasta llegar al plano léxico.

En relación con el proceso de comunicación lingüística, se expondrá cómo una serie de factores implicados condicionan el acto comunicativo y son capaces de conferir a la comunicación un carácter especializado.

A continuación, de los modelos existentes que conciernen a la relación entre lengua general y lenguajes de especialidad, se adoptará la postura defendida por Cabré (1993) y se describirá su modelo. Los lenguajes especializados se conciben como subconjuntos del lenguaje global, diferenciados de la lengua general por aspectos pragmáticos (temática, usuarios y situación comunicativa). Asimismo, los lenguajes especializados comparten una serie de rasgos lingüísticos (sintáctico-gramaticales, léxicos, funcionales y textuales) que les proporcionan unidad entre ellos. Dicha unidad se hace patente en el tipo de discursos especializados, puesto que participan de un mismo aspecto semántico global y formal, e incorporan en el nivel léxico una terminología específica.

Finalmente, se introducirá el léxico de los lenguajes de especialidad y su relación con el léxico de la lengua general. Las unidades léxicas específicas de cada lenguaje de especialidad constituyen una terminología, que distingue a un lenguaje de especialidad de la lengua común y del resto de lenguajes de especialidad.

2.1 Comunicación general y comunicación especializada.

La labor de la lingüística es la descripción y el análisis del lenguaje como objeto de estudio. Describir el lenguaje implica la explicación detallada de todos sus aspectos fundamentales: de su estructura, del uso que los hablantes hacen de él (actuación) y de la forma como se adquiere (aprendizaje o adquisición).

La lingüística teórica se centra fundamentalmente en la descripción de la competencia lingüística, entendida como “un conjunto de conocimientos lingüísticos

que el hablante de cualquier lengua tiene interiorizados, y gracias a los cuales puede construir y entender las oraciones de su lengua” (Cabré, 1993:73). Sin embargo, describir el lenguaje supone dar cuenta no sólo de la competencia del hablante en general, sino también del uso que un hablante concreto hace de la lengua para comunicarse en determinadas situaciones y en la sociedad a la que pertenece.

La comunicación lingüística consiste en “la *interacción* lingüística que, regida por máximas y reglas, ejercen recíprocamente los interlocutores que *negocian* en torno al contenido de un tema o de una materia, con el fin de intercambiar información, expresar opiniones y sentimientos, solicitar ayuda, etc.” (Alcaraz, 2000:85).

Esta concepción de comunicación procede del modelo tradicional propuesto por Jakobson (1960), donde se recogen los principales elementos que intervienen en el proceso comunicativo y que se muestran en la siguiente figura:

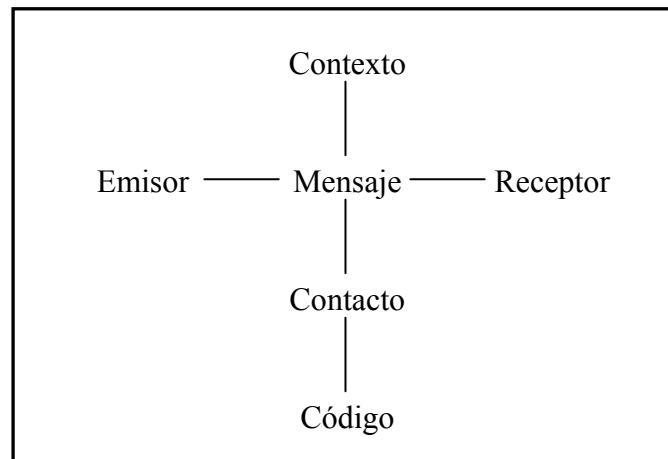


Figura 2.1 Modelo de comunicación según Jakobson (1960).

Aunque el modelo de Jakobson ha gozado de gran predicamento y aún se utiliza en análisis lingüísticos, Alcaraz (2000:86) lo considera insuficiente, “porque en él la comunicación está muy basada en la lengua como código o *langue* y, como tal, se nos presenta como una actividad verbal muy aséptica que no tiene en cuenta las bases cognitivas y sociológicas del lenguaje”. Por tales razones, Alcaraz propugna la inclusión de un modelo de comunicación de base sociológica y toma en consideración el modelo de Hymes (1972), quien sostiene que la comunicación, entendida como la interacción lingüística entre humanos, consta de tres planos:

- a) Situaciones de habla (marcos espacio-temporales donde ocurre la interacción);
- b) acontecimientos de habla (actividades que tienen lugar en la situación de habla y que están gobernadas por normas, convenciones o reglas); y,
- c) actos de habla (acciones que realizan los hablantes para comunicarse con sus interlocutores en un determinado contexto).

La lingüística aplicada concibe el lenguaje precisamente en su dimensión social, como estructura y como herramienta de comunicación, como sistema y como materia de resolución de necesidades comunicativas e informativas que la sociedad tiene planteadas. El lenguaje, visto desde la actuación, se concibe como un sistema, integrado a su vez por varios subsistemas interrelacionados (Cabré, 1993:75). Los hablantes están condicionados por tres variables que dan lugar a los sistemas dialectales naturales: la procedencia geográfica (geolectos o dialectos geográficos), el grupo social (sociolectos o dialectos sociales) y la generación (cronolectos o dialectos históricos).

Igualmente, un hablante se puede encontrar en diferentes situaciones comunicativas que le exigen adecuar su expresión en consonancia, generando las variedades funcionales o registros. Los registros pueden materializarse según varios criterios (Cabré, 1993:77):

- a) el canal utilizado para transmitir la información (oral/escrito)
- b) el tema de que se trata (general/especializado)
- c) el propósito comunicativo o tenor funcional (informar/evaluar/influir/argumentar)
- d) el grado de formalidad entre los interlocutores y el nivel de abstracción de la información (formal/informal; más abstracto/menos abstracto).

De este modo, el lenguaje se concibe como un sistema de sistemas y como un sistema heterogéneo en función de sus variedades dialectales y funcionales. En la presente tesis, el foco de estudio se dirige hacia la variedad del lenguaje determinada por la especialización temática de una comunidad discursiva. El proceso de comunicación entre los hablantes que integran dicha comunidad, y sobre temas especializados, es muy similar al proceso de la comunicación general. Ésta adquiere la calidad de comunicación especializada en función de las restricciones que se imponen a los componentes del acto comunicativo (Cabré, 1993:105): Los interlocutores tipo son especialistas, en mayor o menor grado, de una materia (aunque también se produce comunicación especializada entre un emisor especialista y un receptor lego). Estos especialistas se comunican compartiendo un mundo de referencia limitado al ámbito propio de la especialidad, que se conceptualiza más formalmente que el mundo expresado por el lenguaje general. El sistema de comunicación utilizado en estas situaciones, se caracteriza por incorporar una terminología específica y compartir con el lenguaje general, la sintaxis, morfología y parte del léxico, dando lugar a un subsistema constituyente del sistema global de la lengua. Por último, en la clase de textos producidos predomina la función referencial establecida por Jakobson, relacionada esencialmente con textos informativos y descriptivos.

2.2 Definición de lenguajes de especialidad.

La relación existente entre la lengua general y los lenguajes de especialidad se evidencia en la argumentación del modelo del lenguaje defendido por Cabré. La autora concibe la lengua, dentro del paradigma de la pragmática, como “un conjunto diverso de subcódigos que el hablante utiliza en función de sus modalidades dialectales, o que selecciona en función de sus necesidades expresivas y de las características de cada situación comunicativa” (Cabré, 1993:128). Al mismo tiempo, un conjunto de unidades y reglas, conocidas por todos sus hablantes, rigen el conjunto total de la lengua. De acuerdo con esta concepción, define lenguaje de especialidad frente a lengua general de la siguiente manera (Cabré, 1993:128):

El conjunto de reglas, unidades y restricciones que forman parte del conocimiento de la mayoría de hablantes de una lengua constituye la llamada lengua común o general, que representa un subconjunto de la lengua entendida en sentido global. Las unidades de la lengua común se utilizan en situaciones que pueden calificarse como ‘no marcadas’.

En contraste, hablamos de lenguaje de especialidad (o de lenguajes especializados), para hacer referencia al conjunto de subcódigos –parcialmente coincidentes con el subcódigo de la lengua común – caracterizados en virtud de unas peculiaridades ‘especiales’, esto es, propias y específicas de cada uno de ellos, como pueden ser la temática, el tipo de interlocutores, la situación comunicativa, la intención del hablante, el medio en que se produce intercambio comunicativo, el tipo de intercambio, etc. Las situaciones en que se utilizan los lenguajes de especialidad se pueden considerar, en este sentido, ‘marcadas’.

En este modelo lingüístico, los lenguajes de especialidad se encuentran en relación de inclusión respecto de la lengua global y en relación de intersección respecto de la lengua común o general y el resto de subsistemas; entre ellos comparten características y mantienen una relación de trasvase constante de unidades y convenciones, como ilustra la Figura 2.2 adaptada de Cabré (1993:140).

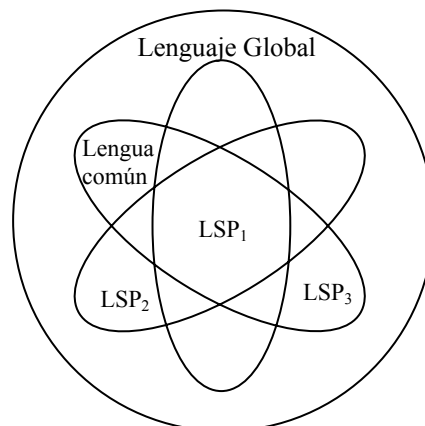


Figura 2.2 Relación entre lengua global, lengua común y lenguajes de especialidad.

Desde la perspectiva de la pragmática, que contempla el uso que el hablante hace de la lengua y no sólo el sistema de reglas o de unidades, un lenguaje de especialidad es “un conjunto de posibilidades determinadas por los elementos que intervienen en cada acto de comunicación: los interlocutores (emisores y destinatarios, con todo el conjunto de características que les son propias), las circunstancias comunicativas, y los propósitos o intenciones que se propone el acto de comunicación” (Cabré, 1993:129).

Esta posición sostiene la imposibilidad de definir los lenguajes de especialidad en términos estrictamente lingüísticos, puesto que son los rasgos pragmáticos o extralingüísticos los que diferencian los lenguajes de especialidad de la lengua general (Sager, Dungworth y McDonald, 1980; Cabré, 1993). Los criterios pragmáticos fundamentales que caracterizan de forma distintiva a los lenguajes de especialidad son la temática, los usuarios y la situación comunicativa:

1. Temática.

Las temáticas especializadas no forman parte del conocimiento general de los hablantes de una lengua y por tanto, han sido objeto de un aprendizaje especializado.

2. Usuarios.

Los usuarios tipo son un subgrupo restringido dentro de la comunidad hablante global, definido por la profesión o la especialización adquirida por aprendizaje: los especialistas. Éstos son productores de lenguaje especializado, cuya audiencia oscila desde receptores también especialistas, hasta aprendices de la materia en cuestión.

3. Situación comunicativa.

Las situaciones comunicativas condicionan el carácter especializado de un subcódigo, puesto que son de tipo formal, normalmente reguladas por criterios profesionales o científicos.

No obstante, estos subconjuntos especializados por la temática, por los usuarios y por las situaciones comunicativas en que se actualizan, también presentan una serie de rasgos lingüísticos que les confieren unidad:

1. Rasgos sintáctico-gramaticales.

Determinadas categorías y estructuras ocurren con mayor frecuencia en los lenguajes especializados que en la lengua general.

2. Rasgos léxicos.

Uso de una terminología concreta que puede presentar distintos niveles de especialización, según el tipo de materia y su nivel de abstracción.

3. Rasgos funcionales.

Predominio de la función referencial del lenguaje. La función básica de los lenguajes de especialidad es la transmisión de información de manera objetiva y precisa. Para ello, especialistas y profesionales recurren a la correspondiente terminología, utilizada para denominar los conceptos del área especializada.

4. Rasgos textuales.

El lenguaje especializado se materializa en determinados tipos de textos y documentos, caracterizados por una configuración global específica y un estilo de carácter regular. El estilo varía en función del grado de abstracción en que se presenta la temática, o el tipo de situación comunicativa a que debe dar respuesta el usuario.

Asimismo, en las comunicaciones o discursos (orales y escritos) especializados elaborados en el seno de los lenguajes de especialidad intervienen una serie de aspectos que les dotan de un carácter unitario. Esta unidad se hace posible gracias a tres tipos de elementos (Cabré, 1993:148):

El aspecto semántico global: se trata de textos concisos (que tienden a ser poco redundantes), precisos (que tienden a no presentar ambigüedad) y despersonalizados (esto es, poco emotivos).

Los elementos que componen la frase: el léxico es el nivel más importante en esta clase de textos, y, dentro del léxico, las nominalizaciones y las formas nominales (por encima de las verbales y de las adjetivas), que desempeñan un papel prioritario tanto cualitativa como cuantitativamente.

El aspecto formal del discurso: los lenguajes de especialidad priorizan la forma escrita por encima de la oral y se distinguen por integrar otros sistemas semióticos en el texto.

El conjunto de características pragmáticas y lingüísticas que comparten los lenguajes de especialidad permiten considerarlos como un subconjunto unificado dentro de la lengua global (Rondeau, 1983; Varantola, 1986; Cabré, 1993).

2.3 El léxico de los lenguajes de especialidad.

En el apartado anterior se hace constar que el nivel léxico es el más importante en el discurso especializado. La razón que justifica tal afirmación se basa en la presencia, dentro del nivel léxico, de la terminología específica propia de la temática correspondiente; terminología entendida como el conjunto de términos de cada materia específica. Numerosos autores coinciden en la importancia de la terminología como elemento que distingue no sólo los lenguajes de especialidad de la lengua común, sino

también los diferentes lenguajes de especialidad entre sí (Sager; 1980; Rondeau, 1983; Cabré, 1993; Alcaraz, 2000).

El nivel léxico está compuesto por un conjunto de lexemas o unidades léxicas de diferentes tipos. Al realizar una comparación entre la selección léxica de un texto especializado y un texto general, se pueden establecer tres grupos de lexemas (Cabré, 1993:152):

- a) El léxico común a ambos textos.
- b) El léxico específico de cada texto, que formaría parte de una capa fronteriza entre la lengua común y la especializada.
- c) El léxico claramente específico del texto especializado, peculiar en virtud de su especialidad y que difícilmente aparecería en textos de la lengua común.

En los textos especializados se hallan unidades léxicas funcionales y de contenido propias del lenguaje común, más las unidades léxicas especializadas propias de la temática, denominadas términos o unidades terminológicas. Al considerar la terminología como constituyente del sistema léxico, que a su vez forma parte del sistema gramatical general, los términos son analizables lingüísticamente como unidades sígnicas, desde tres perspectivas diferentes (Cabré, 1993:171): formal (la denominación), semántica (el concepto) y funcional (la categoría y la distribución).

- a) Perspectiva formal.

Formalmente, los términos son “conjuntos fonológicos susceptibles de ser articulados fonéticamente – y de ser representados gráficamente – con una estructura interna constituida por morfemas”.

- b) Perspectiva semántica.

Semánticamente, los términos son “unidades de referencia a una realidad, y por lo tanto están dotados de un significado que puede describirse como un conjunto de rasgos distintivos”.

- c) Perspectiva funcional.

Funcionalmente, los términos son “unidades distribucionales que requieren un entorno lingüístico determinado, y que en el discurso se encuentran frecuentemente combinados con otros términos específicos (la fraseología)”.

Los términos, en su condición de signos lingüísticos, son unidades de doble cara. Por un lado, presentan la expresión, que se manifiesta por medio de la denominación; y por otro lado, el contenido, que representa la noción o concepto a que se refiere la denominación. La norma ISO/704 de 1987 define concepto o noción como “construcciones mentales que sirven para clasificar los objetos individuales del mundo exterior o interior a través de un proceso de abstracción más o menos arbitrario”.

Conceptos y nociones confieren a los términos la propiedad de referencia a través de las denominaciones. La terminología representa la estructura conceptual y nocional que articula la realidad exterior o interior de una materia especializada.

Cada materia está provista de una terminología propia, cuyos términos, a diferencia de las palabras del léxico común, son “unidades unívocas (la relación entre forma y concepto es única) y monorreferenciales (un término sólo designa un objeto)” (Cabré, 1993:213). De acuerdo con la autora, tal afirmación es más fácil de sostener en la teoría que en la práctica porque, en primer lugar, la terminología no queda exenta de los fenómenos de la sinonimia, homonimia y polisemia; y en segundo lugar, porque resulta difícil demarcar los límites que separan las terminologías de campos afines, al igual que acotar de forma precisa léxico común y terminología. Esto es debido principalmente a las características que comparten los lenguajes de especialidad con la lengua general, y al constante trasvase de unidades léxicas entre los subsistemas. En consecuencia, contenidos y expresiones especializadas de una materia pueden pasar a ser comunes o especializadas en otro ámbito. Por esta razón, suele aceptarse la distinción entre “pertenecer a un campo especializado” y “usarse en un campo especializado” (Cabré, 1993:168). Aunque, en principio, un término pertenece sólo a un dominio especializado, su uso se puede extender a otros campos de especialidad. No obstante, los términos se definen y adquieren su valor semántico dentro del área a la que pertenecen, cuando se activan sus rasgos pragmáticos.

CAPÍTULO 3

EL INGLÉS CON FINES ESPECÍFICOS: EL INGLÉS PROFESIONAL Y ACADÉMICO

3.1 Inicios y definición.

La necesidad de aprender el lenguaje relacionado con el comercio, es decir, de instruirse en la lengua con un propósito específico, se puede situar en el siglo XVI en el caso del inglés, aunque el estudio de las lenguas con una finalidad específica se remonta a tiempos tan lejanos como la Antigüedad Clásica (Dudley-Evans y St. John, 1998:1). No obstante, el inicio del movimiento *English for Specific Purposes (ESP)* se encuentra firmemente ubicado en la segunda mitad del siglo XX. Tras la Segunda Guerra Mundial, se producen una serie de factores que conducen a la identificación del *ESP* como área de conocimiento, dentro de la enseñanza de lenguas extranjeras o segundas lenguas.

El término *ESP* se acuña en inglés por oposición al concepto de *English for General Purposes (EGP)*: “*In teaching English as a foreign language we can speak of English of general purposes (EGP) and English for special purposes (ESP), which represents a division of convenience for designing syllabuses and course outlines*” (Sager, Dungworth y McDonald, 1980).

ESP, traducido al español como inglés con fines específicos (IFE), ha ido adoptando distintas variaciones dependiendo de la “especificidad” que tratara. De este modo, surgen los términos *English for Academic Purposes (EAP)*, *English for Occupational Purposes (EOP)*, *English for Business (EB)*, etc.

El inglés propio de las telecomunicaciones constituye una de las ramificaciones que abarca la variedad del *English for Science and Technology (EST)* o inglés científico-técnico. En *Episodes in ESP*, Swales (1985) consigue identificar la primera publicación sobre las características distintivas de la prosa científica: *Some measurable characteristics of modern scientific prose*, escrito por Barber en 1962. Esta descripción es completada más adelante por Herbert (1965), Lackstrom (1975) y una sucesión de doce estudios más, recopilados en el mismo volumen, que intentan aportar una visión global de la enseñanza del inglés científico-técnico hasta 1985.

Una definición general del lenguaje de especialidad evoca el lenguaje específico utilizado por un grupo de hablantes para comunicarse sobre un tema concreto. De acuerdo con Widdowson (1998), cualquier uso que se haga del lenguaje siempre será “específico”, puesto que siempre servirá para un fin concreto. En esta misma línea, Alcaraz (2000:15) afirma que existen tantas lenguas de especialidad como actividades humanas; considerando lengua de especialidad como “el lenguaje específico que utilizan algunos profesionales y especialistas para transmitir información y para negociar los términos, conceptos y conocimientos de una determinada área de conocimientos”. Estas lenguas de especialidad han formado un núcleo de subvariedades al que Alcaraz denomina “inglés profesional y académico (IPA)”, refiriéndose a la lengua inglesa de especialidad, utilizada en un entorno académico y en la correspondiente carrera profesional.

En la segunda mitad de la década de los 90, Dudley-Evans intenta compendiar la actividad que se había desarrollado hasta la fecha bajo el abanico del ESP. En el artículo titulado *An Overview of ESP in the 1990s* (en Orr, 1997), Dudley-Evans identifica tres aspectos principales que distinguen al ESP de cualquier otra rama de la enseñanza de la lengua inglesa: “*needs analysis, the analysis of genres and language related to these needs, and the use of the methodology of the disciplines or professions it is serving for at least some of the time in materials in the classroom*”. A continuación, introduce una definición del término en función de características absolutas y variables, que desarrolla un año después en *Developments in ESP* junto a St. John:

- *Absolute characteristics.*
 - *ESP is designed to meet specific needs of the learner.*
 - *ESP makes use of the underlying methodology and activities of the disciplines that it serves.*
 - *ESP is centered on the language appropriate to these activities in terms of grammar, lexis, register, study skills, discourse and genre.*
- *Variable characteristics.*
 - *ESP may be related to or designed for specific disciplines.*
 - *ESP may use, in specific teaching situations, a different methodology from that of General English.*

Asimismo, Dudley-Evans completa esta definición con una serie de indicaciones adicionales sobre los destinatarios de los cursos de ESP y las distintas funciones que ha de desempeñar el docente. Los cursos de ESP se pueden diseñar para adultos en la enseñanza universitaria o en situación laboral activa y para la enseñanza secundaria; generalmente están dirigidos a alumnos con un nivel intermedio o avanzado de la

lengua. Respecto al docente, éste debe estar preparado para adoptar el papel de profesor, diseñador del curso y materiales, colaborador, investigador y evaluador.

Otros autores han introducido pequeños matices a esta definición, pero siempre tomando como referencia a Dudley-Evans y St. John. Los rasgos distintivos del ESP llevan a caracterizarlo como un enfoque de la enseñanza, que dirige su atención especialmente hacia la calidad de los materiales didácticos y el resultado del aprendizaje, más que hacia una teoría del propio aprendizaje.

En las últimas décadas, el inglés se ha consolidado como la lengua de comunicación internacional en el comercio y los negocios. Además, en el ámbito científico y académico, las publicaciones de mayor prestigio y difusión mundial se realizan en la lengua inglesa. Según Alcaraz (2000:16) destacan tres razones que justifican el auge que en los últimos años ha experimentado el IPA (término acuñado por el propio autor):

- El desarrollo espectacular de la ciencia y de la tecnología,
- la mundialización o globalización de la economía, y
- el intercambio de estudiantes y profesores entre universidades de distintos países.

En la actualidad, es preferible utilizar el término “inglés profesional y académico” tal y como lo define Alcaraz, porque es más concreto que el de “inglés para fines específicos”, su uso en español se ha extendido ampliamente y así se ha consolidado.

3.2 Evolución del IPA.

El IPA surge como respuesta a una necesidad lingüística generada tras una serie de cambios producidos en el mundo de la economía y la educación. El dominio económico que los Estados Unidos comienzan a ejercer tras la Segunda Guerra Mundial provoca la necesidad de comunicarse en inglés, principalmente en el mundo de los negocios, la ciencia y la tecnología. Todo esto se ve reforzado por el aumento del número de alumnos internacionales que estudian en el Reino Unido, los EE.UU. y Australia.

Desde el reconocimiento del IPA como disciplina, éste ha ido evolucionando paralelamente a la enseñanza de la lengua inglesa en general y la lingüística aplicada, caracterizándose por las mismas tendencias y marcos teóricos. De la misma forma que la teoría del aprendizaje de la lengua condiciona los métodos de enseñanza del IPA, la visión del texto adoptada en cada etapa determina toda su práctica.

Una vez que el IPA se consolida, se van publicando diversos trabajos que recopilan las ideas generadas y la actividad producida en el transcurso de los años. Swales (1985), Kennedy y Bolitho (1984), Hutchinson y Waters (1987), Robinson (1991), Dudley-

Evans y St. John (1998), Alcaraz Varó (2000) y Swales (2000) muestran el advenimiento, los avances y la evolución del inglés profesional y académico. De acuerdo con varios de estos autores, Nelson (2000) ilustra en la siguiente tabla las diferentes etapas de desarrollo que se han sucedido en la historia del IPA. La última fila se ha añadido a la tabla original para mostrar la propia perspectiva de Nelson sobre la visión general de la evolución del IPA:

Robinson 1980	1. Register analysis	2. Discourse analysis and the communicative approach	3. Student motivation and analysis of needs			
Coffey 1984	1. Register analysis	2. a) Discourse analysis b) Notional/functional/communicative approach	3. Needs analysis			
Hutchinson & Waters 1987	1. Register analysis	2. Rhetorical/ Discourse Analysis	3. Target situation analysis	4. Skills and strategies	5. Learning-centred approach	
Johns 1991	1. Register analysis	2. Functional/ Discourse approach	3. Target situational analysis			4. Learning-centred approach
West 1997	1. Authenticity: a) skill based b) skills and strategies 2. Research: a) Register analysis b) Newspeak	3. Text: a) Discourse analysis b) Genre analysis	4. Need: a) Target situational analysis b) Pedagogic needs analysis: deficiency, strategy and means analysis			5. Learning-centred approach
Dudley-Evans & St. John 1998	1. Register analysis	2. Rhetorical/ Discourse Analysis	3. Analysis of study skills	4. Analysis of learning needs	5. No real dominating approach	
Nelson 2000	1. Register analysis	2. Discourse or Rhetorical Analysis. Genre analysis as later development	3. Needs analysis a) Target situation analysis b) Communicative syllabus design, means analysis, deficiency analysis, language audit	4. Skills and strategies	5. The learning-centred approach	6. ESP Today: eclectic period

Tabla 3.1 Etapas en la evolución del IPA.

Es importante señalar que, aunque estas etapas se suceden cronológicamente y se distinguen claramente, también se solapan. Los diferentes enfoques adoptados en el IPA han ido evolucionando y avanzando en el transcurso del tiempo, de modo que en la actualidad, es posible identificar actividad diligente relacionada con cada uno de ellos.

A continuación, se presentan las tendencias lingüísticas más influyentes en el desarrollo del IPA hasta la fecha.

3.2.1 Análisis del registro.

La mayoría de los autores coinciden en considerar el análisis del registro como la etapa inicial del IPA. La primera vez que se utiliza el concepto de análisis del registro es en la obra de Halliday, McIntosh y Stevens (1964): *The Linguistic Sciences and Language Teaching*. Se parte de la idea de que los textos (orales y escritos) utilizados en determinados contextos y entornos, como el académico, de negocios, tecnológico, etc., muestran rasgos especiales que los caracterizan y diferencian de los textos del lenguaje general. El objetivo perseguido consiste en identificar las características gramaticales y léxicas de estos lenguajes especiales o registros, y para ello se valen de métodos estadísticos. Los trabajos más influyentes en este periodo fueron, entre otros, *Some Measurable Characteristics of Modern Scientific Prose* de Barber (1962) y *Writing Scientific English* de Swales (1971).

Johansson (1975) manifiesta la necesidad de diseñar un corpus lingüístico computarizado para satisfacer las necesidades descriptivas de los registros lingüísticos e identificar sus rasgos léxico-gramaticales. El análisis del registro centra su estudio en los elementos estructurales de la oración y los léxico-gramaticales más recurrentes. Sin embargo, este aspecto recibe las principales críticas por detenerse en un análisis descriptivo y no explicativo. Además, al limitarse al nivel oracional, no se aporta información sobre otras características textuales interoracionales. De este modo, los cambios que suceden al análisis del registro se dirigen hacia un nivel extraoracional, motivados por el interés de averiguar los mecanismos lingüísticos que consiguen la cohesión y coherencia en el discurso.

3.2.2 Análisis retórico o análisis del discurso.

Mientras que el análisis del registro centra su estudio en el nivel oracional, la segunda etapa de la evolución del IPA orienta su foco de atención hacia el nivel supraoracional: cómo las oraciones se combinan en el discurso para producir significado. Selinker y Widdowson son figuras destacables como precursores de esta tendencia.

Desde la perspectiva del análisis retórico se pretenden identificar los patrones o esquemas de organización del texto y especificar los medios lingüísticos con los que se construyen tales patrones. Esta etapa se encuentra estrechamente ligada al enfoque funcional. El discurso se considera como una sucesión de actos retóricos (definiciones, clasificaciones) y nociones relacionadas (la causa, la consecuencia, el contraste).

Aunque se sustituye el estudio de las estructuras gramaticales por las funciones del lenguaje, aún no se le otorga al contexto la correspondiente relevancia.

Widdowson y Allen (1974) advierten la relación entre forma y uso de la lengua, y apuestan por un estudio del lenguaje que desarrolle un conocimiento sobre el uso de las oraciones en cada acto comunicativo diferente; introducen así un sílabo o programa nocional/funcional y defienden un enfoque basado en la comunicación. De aquí surgen los primeros atisbos del interés por las destrezas lingüísticas y por las necesidades de los estudiantes, sometidas a un examen completo en la siguiente etapa.

3.2.3 Análisis de la situación término (*Target situation analysis*).

En este periodo, la atención se centra no sólo en el texto sino también en el análisis de la situación término y en las necesidades del alumno. El propósito es identificar la situación concreta en la que se va a utilizar la lengua y analizar las características lingüísticas de la situación. Tales características suponen una necesidad de aprendizaje para el alumno.

Este concepto conduce a la elaboración de varios estudios sobre el análisis de necesidades. Munby (1978) desarrolla el modelo de análisis más exhaustivo: *The Communication Needs Processor*, por medio del cual se identifican las microdestrezas lingüísticas y las funciones semánticas, que se plasman a su vez en formas lingüísticas concretas. No obstante, este modelo recibe numerosas críticas por las dificultades que conlleva su implementación y por no involucrar a los alumnos en el proceso. Como consecuencia, surgen otros modelos que intentan complementar la propuesta de Munby, definiendo las necesidades en términos de *means*, *lacks* y *learning strategies* (Nelson, 2000): *Means analysis*, *deficiency analysis* y *language audit*.

3.2.4 Destrezas y estrategias.

En la década de los 80 tiene lugar un cambio de perspectiva que marca una nueva fase en la evolución del IPA. Se produce un intento por sobrepasar las características lingüísticas superficiales de la situación término y considerar el proceso cognitivo que subyace en el uso del lenguaje, centrándose a la vez en el desarrollo de las destrezas lingüísticas. Desde el periodo del análisis del discurso, el entrenamiento de las destrezas se atenía a la comprensión y expresión escrita. Ahora, la comprensión y expresión oral adquieren una importancia similar.

La idea principal de este enfoque se sustenta en el hecho de que todas las lenguas comparten procesos de razonamiento e interpretación, que permiten extraer significado del discurso. Por lo tanto, el objeto de estudio se fija en “*the underlying interpretive strategies, which enable the learner to cope with the surface forms*”

(Hutchinson y Waters, 1987:13). De este modo, no sería necesario distinguir diferentes registros ya que tales procesos subyacentes son comunes a todos los registros.

Aunque no existe una figura dominante en este enfoque, la mayor parte de los trabajos significativos realizados se encuentran enmarcados dentro de proyectos como el *National ESP Project* en Brasil y el *University of Malaya ESP Project*. Ambos pretenden cubrir unas necesidades de aprendizaje especiales, donde la docencia se imparte en la lengua materna, pero los estudiantes necesitan leer textos especializados disponibles sólo en inglés.

3.2.5 Enfoque centrado en el aprendizaje (*Learning-centred approach*).

Hutchinson y Waters, pioneros en la aplicación del *Learning-centred approach* al IPA, manifiestan que los principios característicos de las etapas precedentes se pueden resumir en tres variables clave: el concepto de necesidad, la concepción del lenguaje y la concepción del aprendizaje. No obstante, puntualizan que estos enfoques están basados solamente en descripciones del uso de la lengua y que no se ha otorgado la importancia suficiente al proceso del aprendizaje.

Our concern is with language learning. We cannot simply assume that describing and exemplifying what people do with language will enable someone to learn it. [...] A truly valid approach to ESP must be based on an understanding of the processes of language learning (Hutchinson y Waters, 1987:14).

Desde esta perspectiva, los autores consideran fundamental descubrir cómo se adquiere la competencia lingüística. La información obtenida del análisis de la situación término no es ignorada por completo, sino que se relega a una segunda posición y se desarrolla el sílabo de forma consecuyente.

3.2.6 Análisis del Género.

Siguiendo un orden cronológico, Nelson (2000) argumenta que la inclinación actual en el IPA se caracteriza por un enfoque ecléctico, donde todas las tendencias principales se encuentran disponibles y son practicables para todos los profesionales. Sin embargo, existe una corriente dominante que deriva del análisis del discurso y que se ha asentado firmemente en el IPA en los últimos años: el análisis del género.

Mientras que el análisis del discurso se centra en identificar las funciones que componen el texto, en el análisis del género, el contexto comunicativo donde se produce el discurso adquiere una gran importancia. En *Genre Analysis: English in Academic and Research Settings* (Swales, 1990:58) Swales, impulsor incuestionable de este enfoque, define y desarrolla el concepto de género:

A genre comprises a class of communicative events, the members of which share some sets of communicative purposes. These purposes are recognized by the expert members of the patent discourse community, and thereby constitute the rationale for the genre. This rationale shapes the schematic structure of the discourse and influences and constrains choice of content and style. Communicative purpose is both a privileged criterion and one that operates to keep the scope of a genre as here conceived narrowly focused on comparable rhetorical action. In addition to purpose, exemplars of a genre exhibit various patterns of similarity in terms of structure, style, content and intended audience.

El análisis del género estudia las formas del discurso producidas por los hablantes de determinadas comunidades discursivas y examina cómo los rasgos léxico-gramaticales forman parte de las convenciones lingüísticas y socioculturales de los textos. De acuerdo con el planteamiento de Swales, el género se equipara con un acontecimiento comunicativo. Sin embargo, otros autores conciben el género como un tipo de texto seleccionado según su objetivo comunicativo. Alcaraz (2000:133) reconcilia estas dos concepciones afirmando que “no son incompatibles sino más bien complementarias” y define género como: “el conjunto de textos, escritos u orales, del mundo profesional y académico, que se ajustan a una serie de convenciones formales y estilísticas”.

La mayor aportación del análisis del género es el esquema propuesto por Swales respecto a la redacción del artículo de investigación y que denomina *Creating a Research Space Model* (CARS model, concebido en 1981 y remodelado en 1990). Swales capta las formas en las que los escritores académicos presentan y justifican su propia contribución a la investigación, de modo que propone seguir una serie de movimientos o *moves* compuestos por diferentes pasos o *steps* cada uno, para redactar la introducción de un artículo de investigación. Aunque este modelo es aplicable a todas las disciplinas académicas, Swales (1990) advierte que se pueden seguir pasos distintos en cada una de ellas. Tras diversas investigaciones, se ha concluido que existen diferencias muy marcadas en la organización de los textos entre culturas diferentes e incluso entre disciplinas diferentes dentro de la misma cultura.

Otros autores han trasladado el *CARS model* al resto de las partes del artículo de investigación con la intención de caracterizar la macroestructura completa: Salager Meyer (1990) y Weissberg (1990) lo aplican a los *abstracts*; Wood (1982), al método; Brett (1994) y Williams (1999), a los resultados; Belanger (1982) y Dudley-Evans (1994), a la discusión.

Al igual que Swales aplica el análisis del género a los géneros académicos, Bhatia implementa la metodología de este enfoque en el ámbito del discurso profesional: *Analysing Genre: Language Use in Professional Settings* (1993), es la obra considerada como el segundo pilar del análisis del género.

3.2.7 Tendencias actuales y líneas de investigación.

El perfil de la evolución de una disciplina académica se puede trazar siguiendo la literatura publicada en revistas especializadas y las conferencias expuestas en congresos y jornadas, que recogen el trabajo realizado en torno a tal disciplina. En el área del IPA, *English for Specific Purposes journal (ESPj)* es la única revista especializada internacional de reconocido prestigio que cubre todas sus ramas. Por lo tanto, una exploración de los artículos publicados durante la existencia de la revista (desde 1980) aportaría una visión bastante representativa de la evolución del IPA y, asimismo, se podrían vislumbrar los pasos a seguir y las futuras líneas de investigación. Éste es el propósito que persigue Hewings en *A History of ESP through 'English for Specific Purposes'* (2002), estudio basado en tres cuestiones principales: el origen geográfico de los artículos, el tema tratado y las referencias bibliográficas.

Respecto a la primera cuestión, Hewings constata un aumento de artículos originarios de fuera de los Estados Unidos y Reino Unido, posiblemente debido al crecimiento de la actividad del IPA en todo el mundo, y su aceptación como disciplina académica en los departamentos de lengua inglesa o lingüística aplicada en la enseñanza superior. Hewings destaca el crecimiento del IPA en Europa y especialmente en España: “[in] Eastern Europe where the number of recent conferences on ESP seems to reflect growing interest in the areas, and also in Mediterranean countries, particularly Spain, where ESP is becoming a focus of special interest in universities.”

En el transcurso de los volúmenes publicados, se observa una disminución constante en el tratamiento del IPA en términos generales, reflejando una tendencia a la especialización en ramas más específicas; entre ellas, el inglés con fines académicos es la más frecuente y el inglés de los negocios muestra una tendencia al alza.

Hewings identifica siete categorías en las que se puede agrupar el contenido de los artículos: “*These are papers which: (i) analyse written text or spoken discourse; (ii) describe a programme or course or part of a course; (iii) focus on needs analysis or syllabus design; (iv) are concerned with materials or methods; (v) present an argument or discussion; (vi) focus on testing; and (vii) deal with teacher training*”.

La relación de los autores más citados vinculados a estas áreas temáticas aporta una idea de las principales influencias en el IPA. Hewings distingue cuatro periodos en la historia del ESPj e identifica los autores más destacados en cada uno. Sin embargo, tres de ellos ejercen una gran autoridad desde los inicios hasta la actualidad: John Swales, Henry Widdowson y Tony Dudley-Evans. Swales (1985, 2000) es citado por su aportación al inglés científico-técnico y su teoría sobre el análisis del género; Widdowson (1974, 1983, 1998) por su labor desarrollada en el enfoque comunicativo de la enseñanza de lenguas específicas; y Dudley-Evans por su trabajo sobre la

redacción en inglés académico y su contribución al establecimiento del IPA como disciplina.

Otros autores mencionados son Selinker, Lackstrom y Trimble (1972) por sus aportaciones sobre el interlenguaje y el inglés científico técnico; Munby (1978) por su trabajo seminal sobre análisis de necesidades; Mackay y Mountford (1978) y Hutchinson y Waters (1987) como referentes de los libros de texto del lenguaje especializado; Krashen (1987) por la influencia ejercida con su trabajo sobre la adquisición de una segunda lengua; y Halliday (1985) por la lingüística funcional sistémica aplicada al análisis del texto.

En los últimos volúmenes de las revistas agrupadas en el periodo más reciente (1997-2001), Hewings advierte dos novedades. La primera es la aparición de Bhatia y Hyland, que refuerzan el área de la redacción académica y, por otro lado, es notable el número de referencias a Sinclair y sus significativas aportaciones a la lingüística del corpus.

Finalmente, como conclusión de su estudio, Hewings se permite augurar las posibles tendencias que seguirán ejerciendo una gran influencia en el desarrollo del IPA en la siguiente década (2001-2010):

- *Internationalisation: ESP practice will continue to spread geographically.*
- *Specialisation: More specific contexts are being analysed, reflecting the increasing specialisation of ESP courses.*
- *Growth of Business English: The growth of Business English in universities is likely to demand and feed into research on business communication.*
- *Continued influence of genre analysis, corpus analysis and systemic functional linguistics.*
- *The effect of English as an international language: The growing use of English as the means of communication in interaction between non-native English speakers.*

Respecto a la situación actual del IPA en España, Bueno (2003) realiza un estudio similar, tomando como fuente las actas de los congresos celebrados en España sobre lenguas para fines específicos (LFE). El estudio abarca el periodo de tiempo comprendido entre 1985, fecha de publicación de la primera comunicación expuesta en España sobre este tema, hasta finales del año 2002.

Bueno clasifica el contenido de las comunicaciones en siete áreas temáticas, coincidiendo con Hewings en tres de ellas: Diseño curricular y enseñanza de LFE; Análisis del discurso. Registros y géneros; y Análisis de necesidades, materiales y

evaluación. Las áreas restantes, que tienen una mayor presencia en la práctica del IPA en España, son por un lado lexicología, lexicografía y traducción, y por otro lado, la adquisición y aprendizaje del lenguaje junto a la teoría sobre LFE (áreas unidas por la autora). A continuación, contando con un número menor de artículos, el uso de medios audiovisuales y nuevas tecnologías componen un área cada vez más presente en la enseñanza de LFE. En último lugar, la autora recoge bajo el mismo abanico dos áreas distintas debido a su reducida presencia en los artículos: Lingüística del corpus y análisis contrastivo.

3.3 Análisis de necesidades.

La elaboración de un curso de IPA requiere una rigurosa planificación, debiéndose tomar en cuenta numerosos factores, principalmente satisfacer la imperiosa necesidad del alumnado de aprender el lenguaje utilizado en un campo específico. Consecuentemente, habrá que realizar un análisis previo para determinar cuáles son las necesidades reales del alumnado en tres niveles distintos: análisis de la situación en la que se utilizará el lenguaje especializado (*target situation analysis*), análisis de la situación del aprendizaje (*learning situation analysis*) y análisis de la situación de los alumnos (*present situation analysis*). Siguiendo el concepto de análisis de necesidades de Dudley-Evans y St. John (1998), éste debería incluir los siguientes aspectos:

- Información sobre la profesión de los alumnos, actividades en las que utilizan o utilizarán el inglés y qué destrezas necesitan para obtener una comunicación efectiva en cada situación.
- Información personal sobre los alumnos, factores que pueden afectar al aprendizaje, como las razones por las que asisten al curso, sus expectativas y su actitud hacia el inglés.
- Información sobre su competencia en la lengua inglesa, lo que permitirá identificar las carencias de los alumnos.
- Información sobre el entorno en que el curso tendrá lugar.

La información más accesible es la resultante del análisis del entorno y situación donde se impartirá el curso. Se trata de un curso de inglés para alumnos de Ingeniería Telemática y de Telecomunicaciones, tiene una duración de 45 horas por cuatrimestre, repartidas en dos sesiones de hora y media a la semana. Los alumnos eligen libremente asistir a uno o a ambos cuatrimestres, puesto que los contenidos son distintos y complementarios.

Su competencia en la lengua inglesa es muy heterogénea. El hecho de que los alumnos puedan cursar esta asignatura en cualquier año de sus estudios da lugar a una

gran diversidad, tanto en la competencia lingüística y comunicativa como en el conocimiento técnico. Es notable la presencia de alumnos que asisten al curso después de haber estado varios años sin contacto con el inglés, pero que poseen un conocimiento más profundo sobre la materia a la que se dedican.

La información respecto a las necesidades relacionadas con el entorno académico y profesional se ha recabado de diversas fuentes. Los mismos descriptores de la titulación describen el perfil del ingeniero de telecomunicación, en función de las tareas que deberá desempeñar en el ejercicio profesional. Éstas implican la investigación y desarrollo; transferencia de tecnología e innovación; diseño, planificación y producción; operación y mantenimiento; soporte a la comercialización y venta de productos; gestión de proyectos y docencia.

La segunda fuente consultada procede de un estudio sobre el perfil del ingeniero que Europa requiere para el siglo XXI, llevado a cabo por el departamento de Lingüística Aplicada a la Ciencia y la Tecnología de la Universidad Politécnica de Madrid (García, 1994). Las conclusiones obtenidas son igualmente extrapolables y válidas para los alumnos de la UPCT y para todos los estudiantes de ingeniería en general.

De acuerdo con los resultados de este estudio, existe una gran demanda de integración del ingeniero español en el contexto europeo. Esta incorporación se puede producir durante el periodo de formación, la participación en programas de prácticas en empresas extranjeras y en el terreno laboral. Necesariamente, es imprescindible el manejo de un idioma común para hacer posible la comunicación: el inglés es la lengua franca para la comunidad científica, académica y profesional de las nuevas tecnologías.

Una de las pautas de mejora de la enseñanza superior marcadas por la política de la Unión Europea, consiste en el conocimiento de otras lenguas. Aunque se pretende promover de forma igualitaria todas las lenguas europeas, el estudio realizado por Hermans (1997) sobre la política lingüística de la UE mediante la implementación del programa Lingua (actualmente Sócrates y Leonardo), muestra cómo es el inglés el idioma más requerido por parte de profesores y alumnos.

Desde el punto de vista de las empresas y de los propios alumnos encuestados, ha sido posible identificar las áreas que requieren una solución y que claramente describen las necesidades comunicativas más apremiantes, las cuales se pueden clasificar en cuatro apartados:

- Recopilación de información: Los alumnos parecen tener un escaso conocimiento de los materiales disponibles escritos en inglés y de los medios para adquirir esta información relacionada con su especialidad.

- Comprensión y resumen: En general no existe mucha dificultad para comprender textos técnicos en inglés, sin embargo la tarea de resumir las principales características de un artículo y expresar las ideas más importantes supone una gran dificultad.
- Expresión escrita: Los alumnos necesitan aprender las estructuras de ordenación lógica de la información, en el estilo adecuado y el uso apropiado de los ejemplos que ilustran la información.
- Presentaciones orales: La exposición oral de una presentación formal ante una audiencia y la realización de una entrevista oral suponen dos áreas de gran dificultad.

Finalmente, la tercera fuente de información consiste en un estudio realizado a menor escala en la UPCT. Se han recogido las opiniones de los alumnos y se ha realizado un seguimiento posterior, con el fin de evaluar el cambio de impresiones y la evolución de la enseñanza del inglés.

Como resultado de las encuestas distribuidas desde el curso 1999/00, año en el que comienza la docencia de Ingeniería Telemática y de Telecomunicación y con ellas la oferta de cursos del correspondiente inglés específico, los alumnos que deciden cursar inglés técnico lo hacen motivados principalmente por las siguientes razones:

1. Para adquirir vocabulario técnico.
2. Para poder optar a mejores salidas profesionales.
3. Para poder manejar bibliografía de su especialidad en inglés.
4. Para poder desenvolverse en inglés en su profesión.

Los contenidos del curso deben cumplir a la vez las expectativas de los alumnos y la demanda de la sociedad, necesidades que el profesor debe satisfacer mediante el estudio de la lengua y facilitando en el reducido entorno de la clase las herramientas necesarias para alcanzar tal fin. Ante esta situación el profesor tiene acceso a varios recursos: puede hacer uso de su intuición y experiencia para elaborar un programa apropiado seleccionando material del que se encuentra disponible, ya que resulta prácticamente imposible encontrar un libro de texto publicado que concuerde exactamente con las necesidades del grupo específico. Otra opción es extraer el material del entorno textual auténtico mediante una rigurosa selección de textos significativos, es decir, cubrir las necesidades de aprendizaje con el resultado del análisis de los textos y géneros representativos utilizados en el contexto profesional y académico. Hutchinson y Waters (1987) afirman:

Given that the purpose of an ESP course is to enable learners to function adequately in a target situation, that is, the situation in which the learners will use

the language they are learning, then the ESP course design process should proceed by first identifying the target situation and then carrying out a rigorous analysis of the linguistic features of that situation. The identified features will form the syllabus of the ESP course.

Los actos de comunicación que tienen lugar en estas situaciones comunicativas específicas se encuentran categorizados por medio de los géneros que se manifiestan en el aula en forma de textos orales o escritos. Por lo tanto, el propósito principal del curso está enfocado hacia el manejo eficiente del discurso especializado a través del cual se genera, transmite y adquiere el conocimiento. Cuando el comportamiento lingüístico de esta comunidad de hablantes se encuentra registrado y materializado por medio de textos potencialmente analizables, es posible hallar los patrones del uso idiomático de la lengua. De este modo, el texto se convierte en fuente de problemas conceptuales y soluciones didácticas, adquiriendo a la vez un carácter motivador para el alumno, puesto que responde a sus necesidades profesionales concretas y específicas (Alcaraz, 2000).

En el estudio del inglés como lengua extranjera, el material didáctico desempeña un papel crucial en la exposición de los alumnos al lenguaje, convirtiéndose en la única fuente y contacto con la lengua inglesa. Aunque en el caso de estudiantes de ingeniería el contacto aumenta debido a que gran parte de las asignaturas de contenido introducen términos en inglés y hacen referencia a bibliografía en inglés, el lenguaje se utiliza en muchos otros contextos donde los alumnos deberán desenvolverse. Además, el lenguaje de los libros de texto y científicos, aun siendo real, no supone más que una parte de la variedad total. Un corpus lingüístico específico ofrece la oportunidad de enfrentarse a una gran diversidad de muestras auténticas de la lengua, que exhiben el uso de las palabras en situaciones determinadas y que proporcionan la exposición al lenguaje de difícil acceso para los alumnos.

Como muestra auténtica del lenguaje se entiende cualquier texto que no se ha producido con un fin didáctico sino de forma natural. Normalmente estos textos se utilizan en el entorno académico o profesional específico, por lo tanto siempre versarán sobre un tema específico en mayor o menor grado, relacionado con el conocimiento de los alumnos sobre la materia, y mostrando siempre el uso real del lenguaje. Las muestras son exactamente actos reales de comunicación de la comunidad discursiva.

La utilización de textos no adaptados es muy provechosa, puesto que esta fuente indica con mayor certeza el tipo de estructuras sintácticas y léxicas más frecuentes. Partiendo de esta información, se pueden elaborar tareas y actividades más aproximadas a las necesidades reales de los alumnos. Este material específico funciona como portador del contenido, del aspecto lingüístico o de la destreza que se está trabajando realmente en una tarea.

Cabe resaltar una cuestión esencial relacionada con este contenido específico: el conocimiento que el profesor tiene sobre la materia. Los alumnos, generalmente, no esperan que el profesor de inglés posea conocimientos de ingeniería, sino que sepa cómo se utiliza el inglés en ingeniería. Esto requiere del profesor un conocimiento básico sobre la materia. Para que se produzca una comunicación significativa entre docente y discente, debe existir un fondo de conocimientos e interés común, algo que los alumnos también suelen apreciar y valorar.

Teniendo en cuenta las necesidades detectadas y la carencia de material didáctico específico, la compilación y análisis de un corpus lingüístico especializado en telecomunicaciones proporcionaría la clave para diseñar el programa y facilitar material didáctico adecuado.

Los estudios basados en corpus están aportando cada vez más pruebas convincentes para discernir qué elementos del lenguaje es más probable encontrar y que, por lo tanto, requerirán una inversión mayor de tiempo en la enseñanza.

La probabilidad de aparición y la información sobre la distribución de los elementos del lenguaje constituyen indicios de utilidad. Tales marcadores pueden influir en la selección de contenidos, en la secuenciación de la enseñanza y en la importancia atribuida a las diferentes partes del lenguaje.

3.4 Contexto metodológico. El análisis del texto según la tradición británica: Firth, Halliday y Sinclair.

La corriente lingüística británica impulsada por John R. Firth en los años 30, sienta la base metodológica para el análisis del lenguaje en la que se ubica esta tesis. Las ideas de Firth han ido evolucionando y, en la actualidad, su concepción del lenguaje se hace visible principalmente en las figuras de Michael Halliday y John Sinclair. Concretamente, la contribución de Sinclair a la lingüística del corpus ha sido clave para el desarrollo de este trabajo.

Los principios teóricos que fundamentan esta corriente demandan el estudio del uso del lenguaje real (oral y escrito), por medio de corpus de datos y en relación con la función que el lenguaje desempeña en las instituciones sociales. Los principales argumentos giran en torno al análisis del lenguaje tal y como ocurre de forma natural, sin la intervención del analista, y a las implicaciones prácticas del estudio del lenguaje.

En 1957, Firth publicó uno de sus artículos más famosos: “*A synopsis of linguistic theory, 1930-1955*”, el mismo año en el que Chomsky publicó “*Syntactic Structures*”. La comparación de los principios que defienden ambas escuelas muestra diferencias considerables. Stubbs (1996) concentra la visión distintiva del lenguaje de la corriente

británica en nueve principios básicos, cuyo planteamiento desarrolla por oposición a la concepción del lenguaje propuesta por Chomsky.

A continuación, se reflejan los nueve principios que distingue Stubbs (1996:25-44), sustentados por una breve explicación en los siguientes apartados. Asimismo, se dilucida cómo tales principios están relacionados con el presente trabajo:

Principle 1: Linguistics is essentially a social science and an applied science.

Principle 2: Language should be studied in actual, attested, authentic instances of use, not as intuitive, invented, isolated sentences.

Principle 3: The unit of study must be whole texts.

Principle 4: Texts and text types must be studied comparatively across text corpora.

Principle 5: Linguistics is concerned with the study of meaning: form and meaning are inseparable.

Principle 6: There is no boundary between lexis and grammar: lexis and grammar are interdependent.

Principle 7: Much language use is routine.

Principle 8: Language in use transmits the culture.

Principle 9: Saussurian dualisms are misconceived.

Principio 1: La Lingüística es esencialmente una ciencia social y aplicada.

La concepción de la lingüística como disciplina es la primera diferencia entre la escuela chomskiana y la británica. Chomsky concibe la lingüística como una rama de la psicología cognitiva, mientras que Firth, Halliday y Sinclair la consideran una ciencia social aplicada con implicaciones prácticas.

Firth mantiene que el estudio de la lengua es el estudio del significado, y que la naturaleza del lenguaje se encuentra estrechamente relacionada con las necesidades de los hablantes y con la función para la que sirve. Las palabras adquieren significado por el uso que se les otorga dentro del contexto social en el que operan. Por lo tanto, el interés de la lingüística debe ser el significado del acto comunicativo en cada contexto.

El carácter de esta investigación es de naturaleza eminentemente práctica, puesto que los resultados del análisis del lenguaje específico de las telecomunicaciones son aplicables directamente en el aula. Las conclusiones esgrimidas no sólo permitirán elaborar programas didácticos adecuados a las necesidades detectadas, sino que también suponen un gran potencial para la creación de manuales de referencia para profesionales del sector, glosarios de términos especializados, etc.

Principio 2: El lenguaje se debería estudiar en muestras auténticas, acreditadas y reales, y no por medio de frases aisladas, inventadas e intuitivas.

La mayor crítica que reciben los lingüistas chomskianos está relacionada con la forma de obtener los datos lingüísticos y su naturaleza. Estos no consistían en ejemplos observados del lenguaje en uso sino que eran fruto de la introspección de informantes. Los propios lingüistas inventaban frases para ilustrar su teoría, siendo conscientes de que no eran representativas y de que la mayoría no ocurriría nunca en la realidad. Asimismo, confiaban en la intuición del hablante nativo para discernir la corrección gramatical de los ejemplos.

La escuela británica no concibe el estudio del lenguaje a través de frases inventadas: *“Usage cannot be invented, it can only be recorded”* (Sinclair, 1987:xv); ni aisladas y descontextualizadas: *“the complete meaning of a word is always contextual, and no study of meaning apart from a complete context can be taken seriously”* (Firth, 1935 en Stubbs, 1996:30).

Stubbs concluye que la solución idónea sería: *“A large corpus, consisting of at least several million words, searched with computer assistance”*. Un corpus lingüístico registra el uso del lenguaje real y contextualizado, en donde se puede identificar lo central y típico del lenguaje: *“what is central and typical in the language”* (Sinclair, 1991:17); y que permite derivar una teoría procedente de la observación y evidencia de los datos: *“a linguistic description which is not supported by the evidence of the language has no credibility”* (Sinclair, 1991:36).

En el presente trabajo se sigue la tradición británica, optando por el uso de muestras auténticas del lenguaje. Los datos manejados proceden de corpus lingüísticos y en ningún caso son fruto de la introspección o intuición de la analista.

Principio 3: La unidad de estudio debe ser el texto completo.

El fundamento de este principio está basado en la distribución irregular que presentan los elementos lingüísticos. La unidad básica de análisis desde el punto de vista del significado es el texto, entendido siempre dentro del contexto en el que se enmarca. Si no se considera el texto completo, no se está reflejando el comportamiento típico: *“Once it is decided to include a text, then all the instances of all the words constitute the evidence”* (Sinclair, 1991:40). Por consiguiente, en la construcción de un corpus se deberán recoger sólo textos completos, rechazando así los fragmentos de texto que se solían compilar en los primeros proyectos de corpus.

No obstante, este trabajo parte de la palabra como unidad de estudio, siguiendo diversos trabajos realizados (Scott, 1997; Curado, 2000; Nelson, 2000; Almela, 2006; etc.). Firth sostiene que el lenguaje es un todo que se puede analizar a varios niveles. El

nivel léxico forma parte de los niveles básicos de la lengua, considerados como una abstracción taxonómica realizada para analizar un texto. Las unidades léxicas se actualizan como palabras en el texto, consideradas como unidades individuales y como unidades que forman parte de un vocabulario. El léxico se estudia en su co-texto inmediato, enmarcado dentro de un texto completo, que a su vez pertenece a un acto del habla auténtico.

Principio 4: Los textos y tipos de texto se deben estudiar de forma comparada por medio de corpus textuales.

Firth defiende que cualquier elemento en la lengua tiene significado en cuanto que está integrado en un sistema, y que la gramática es un sistema de opciones que el hablante elige para cada acto de habla según las circunstancias. De este modo, el significado de cualquier elemento del lenguaje está relacionado tanto con los elementos que lo acompañan en el mensaje (relaciones sintagmáticas), como con todos los posibles elementos que se pueden usar y que no se han usado en dicho contexto (relaciones paradigmáticas). Consecuentemente, los textos se interpretan en referencia a un fondo de expectativas, y tanto por lo que expresan como por lo que omiten.

Las variedades del lenguaje se manifiestan a través de la variación de la frecuencia de los elementos léxicos y gramaticales en los diferentes tipos de textos. De hecho, la variedad del registro se define normalmente como la variación sistemática de tales frecuencias (Halliday, 1991; Biber, 1998). Por lo tanto, esta visión del lenguaje implica el estudio comparativo de textos; en términos de Sinclair (1965, en Stubbs, 1996):

Any stretch of language has meaning only as a sample of an enormously large body of text; it represents the results of a complicated selection process, and each selection has meaning by virtue of all the other selections which might have been made, but have been rejected.

El análisis de grandes cantidades de textos demuestra que el lenguaje en uso se caracteriza por “*spectacular regularities of patterns with endless variation*” (Sinclair, 1991:4). Estos resultados aportan una sólida base de interpretación para realizar comparaciones empíricas de textos individuales y de diferentes géneros.

En este trabajo se pretende identificar el vocabulario característico del lenguaje de la ingeniería de telecomunicaciones, comparándolo con el lenguaje general por medio de la aplicación de criterios cuantitativos y cualitativos. Para este fin, se ha creado un corpus específico de tal tamaño que permita identificar sus rasgos léxicos distintivos, al contrastarlo con un corpus general.

Principio 5: La Lingüística se ocupa del estudio del significado: forma y significado son inseparables.

Las ideas sostenidas por la escuela británica sobre la interdependencia entre forma y significado se ven reflejadas en Sinclair (1991) al manifestar: “*there is ultimately no distinction between form and meaning. ... [The] meaning affects the structure...*”. Estas ideas son totalmente opuestas a Chomsky (1957), quien afirma: “*Grammar is autonomous and independent of meaning*”.

El hilo conductor que enlaza la tradición británica es el concepto de colocación propuesto por Firth (1957): “*the company a word keeps*”, es decir, una palabra adquiere parte de su significado por las otras palabras que la acompañan. Stubbs añade al respecto que en las colocaciones se producen relaciones sintagmáticas entre palabras como tal, y no entre categorías (Stubbs, 1996:35).

En este trabajo se exploran las relaciones sintagmáticas de las palabras clave halladas en el lenguaje de las telecomunicaciones, por medio del estudio de los denominados colocados. Los colocados son aquellas palabras que ocurren en el entorno de la palabra clave y que contribuyen a describir el significado y uso de dicha palabra clave.

Principio 6: No existen fronteras entre léxico y gramática: léxico y gramática son interdependientes.

Halliday propone el concepto de *lexico-grammar* argumentando que “*lexis is the most delicate syntax*”, aceptando que léxico y gramática se interrelacionan. Sin embargo, esta posición ha ido evolucionando a través de numerosos estudios realizados en el marco de la tradición británica (Francis, 1991; Sinclair, 1992a; Louw, 1993; Stubbs, 1996 y 2001; Hoey, 2005); Nelson (2000:217) lo enuncia de manera clara y sencilla: “*Lexis choses grammar and grammar choses lexis*”. Stubbs (1996:36-40) describe el principio de co-selección y lo ilustra con diversos ejemplos que muestran a su vez la evolución del concepto, extrayendo las siguientes conclusiones:

1. *Any grammatical structure restricts the lexis that occurs in it; and conversely, any lexical item can be specified in terms of the structures in which it occurs.*
2. *Such restrictions are typically not absolute, but clear tendencies: grammar is inherently probabilistic.*
3. *Meaning is not constant across the inflected forms of a lemma.*
4. *Every sense or meaning of a word has its own grammar: each meaning is associated with a distinct formal patterning. Form and meaning are inseparable.*

5. *Words are systematically co-selected: the normal use of language is to select more than one word at a time.*
6. *Since paradigmatic choices are not made independently of position in syntagmatic chain, the relation between paradigmatic and syntagmatic has to be rethought.*
7. *Traditional word-classes and syntactic units also have to be rethought.*
8. *Native speakers have only limited intuitions about such statistical tendencies. Grammar is corpus-driven in the sense that the corpus tells us what the facts are. Some of these facts...cannot be predicted in advance and they certainly cannot be exhaustively documented from intuition.*

Principio 7: Gran parte del lenguaje es rutina.

Este principio se fundamenta en la convicción de que el lenguaje se constituye de elementos léxicos (palabras, colocaciones y agrupaciones léxicas) que se repiten una y otra vez en situaciones habituales (Sinclair, 1991; Nattinger y DeCarrico, 1992). De este modo, los hablantes están condicionados por una serie de posibilidades y restricciones lexico-gramaticales, y cada acto de habla determina en gran medida al siguiente. De nuevo existe una confrontación con la corriente Chomskiana, donde se favorece la creatividad y se considera la rutina del lenguaje de forma negativa.

Este trabajo encauza el estudio de la rutina del lenguaje de las telecomunicaciones, por medio del análisis léxico propio del dominio y de los clusters o agrupaciones léxicas típicas que se producen en el mismo. La identificación y clasificación de las palabras claves, colocados y clusters permiten ilustrar el perfil léxico del lenguaje especializado.

Principio 8: La lengua en uso transmite la cultura.

Cuando Firth propuso el estudio de la transmisión de la cultura a través de los actos de habla, todavía no era posible acceder a corpus lingüísticos y métodos de análisis computacional que facilitaran “*research into the distribution of sociologically important words*” (Firth, 1935, en Stubbs, 1996). Stubbs ejemplifica y argumenta detalladamente cómo se pueden identificar y examinar expresiones fijas y palabras recurrentes, para demostrar cómo la cultura se manifiesta en los patrones léxicos.

La transmisión de la cultura a través del lenguaje no es objeto de estudio en este trabajo.

Principio 9: Los dualismos de Saussure están mal planteados.

Firth, Halliday y Sinclair se oponen directamente a la dicotomía del lenguaje *competence-performance* de Chomsky y, por extensión, a la distinción entre *langue* y

parole de Saussure. *Langue/competence* se concibe como sistemática y como el único objeto de estudio, aunque sea una entidad abstracta y por lo tanto, no observable. Por otro lado, *parole/performance* carece de interés porque se considera idiosincrásica y asistemática. Además, el concepto de *parole/performance* es observable sólo de forma fragmentaria, ya que como un todo es inobservable y, en cualquier caso, no refleja la competencia lingüística innata del hablante-oyente ideal (Stubbs, 1996:44). Un corpus lingüístico permite evaluar muestras auténticas del lenguaje y seleccionar las más típicas: “A complete set of typical instances should exemplify the dominant structural patterns of the language without recourse to abstraction, or indeed to generalization” (Sinclair, 1991:103).

La aportación fundamental de la lingüística del corpus consiste en que el análisis del lenguaje asistido por ordenador, permite acceder a datos que durante años no había sido posible observar. Un corpus no es una simple herramienta de análisis sino un concepto de gran importancia en la teoría lingüística, en términos de Sinclair (1991a:xvii):

Analysis of extended naturally-occurring texts, spoken and written, and, in particular, computer processing of texts have revealed quite unsuspected patterns of language... The big difference has been the availability of data... [The] major novelty was the recording of completely new evidence about how the language is used... [The] contrast exposed between the impressions of language detail noted by people, and the evidence compiled objectively from texts is huge and systematic... The language looks rather different when you look at a lot of it at once.

El marco metodológico adoptado en el presente trabajo permite el estudio de los rasgos típicos y característicos del lenguaje, utiliza muestras reales del lenguaje procedentes de situaciones auténticas, que además se encuentran almacenadas en formato electrónico y disponibles para su análisis estadístico.

4.1 Inicios y evolución de la Lingüística del Corpus.

La Lingüística del Corpus es una disciplina que tiene como objetivo la descripción de la naturaleza, estructura y uso de la lengua, mediante el análisis de muestras reales de uso lingüístico. Esta disciplina se basa en un cuerpo de textos como materia de estudio y fuente de evidencia para la descripción y argumentación lingüística, incorporando a la vez metodologías de descripción que cuantifican la distribución de las unidades lingüísticas.

La Lingüística del Corpus es una disciplina joven. Aunque los primeros trabajos realizados datan de la década de los sesenta, lexicógrafos y lingüistas empezaron a utilizar corpus a finales de los años ochenta; en la enseñanza de lenguas no se introdujo hasta los noventa y su uso directo por los estudiantes de lenguas es aún incipiente.

El desarrollo y difusión de la Lingüística del Corpus es directamente proporcional a la evolución y fácil acceso de los ordenadores. Estos facilitan las tareas de descripción textual y permiten aumentar considerablemente el tamaño de las bases de datos empleadas para el análisis.

No obstante, existe constancia de recopilaciones de textos destinados a fines similares, que datan de años anteriores y se pueden considerar como los antecedentes de la Lingüística del Corpus. Se trata de compilaciones realizadas y analizadas manualmente, sin la asistencia del ordenador. Tribble y Jones (1990) hacen constar que, en el siglo XIII, 500 monjes dirigidos por Hugo de San Charo analizan las concordancias de la Biblia. Utilizando la misma fuente como corpus, Cruden (1736) aporta listados de concordancia de las palabras que juzga más importantes (Kennedy, 1998).

Más adelante, Kaeding (1898) consiguió recopilar once millones de palabras procedentes de las deliberaciones del gobierno alemán. Su intención era extraer información sobre el uso de las palabras y las letras del alemán, y aplicarla en la formación de taquígrafos (Kennedy, 1998).

De nuevo tomando muestras de la Biblia y añadiendo otras fuentes, Knowles (1904) proporcionó un listado de las 353 palabras más comunes de la lengua inglesa. Sin embargo, Thorndike ejerció una influencia mayor con *The Teacher's Word Book* (1932), donde publicó una lista con las 20.000 palabras más frecuentes en inglés. El corpus utilizado para el análisis contenía 4,5 millones de palabras procedentes de clásicos de la literatura inglesa y otras fuentes. Los resultados derivados de este trabajo propiciaron un gran cambio en la enseñanza de la lengua inglesa.

Más tarde, Michael West (1953) elaboró una lista con las 2.000 palabras más frecuentes del inglés escrito. Esta lista se tomó como modelo de referencia en la elaboración de la programación de cursos de inglés para estudiantes extranjeros y se continúa aceptando en la actualidad. *A General Service List of English Words* incluye información detallada sobre las diferentes acepciones de cada palabra y el porcentaje de uso de cada acepción.

Probablemente, el resultado más importante de esta etapa inicial fue el reconocimiento de que una pequeña cantidad de palabras diferentes, generalmente constituye la mitad del total de las ocurrencias de cualquier texto (Kennedy, 1998).

El corpus pre-electrónico que marcó la transición entre las descripciones basadas en corpus no computerizados y el desarrollo de la Lingüística del Corpus actual, fue el proyecto liderado por Randolph Quirk en 1959: *Survey of English Usage* (SEU) Corpus. Su propuesta consistía en recopilar 200 muestras del lenguaje, compuestas por 5.000 palabras cada una y que fueran representativas del inglés británico hablado y escrito. El objetivo era conseguir un corpus de un millón de palabras, que sirviera como base de descripción de la gramática y del uso que los hablantes nativos, adultos y cultos, hacen del inglés británico.

La trascendencia de este corpus estriba en la procedencia oral de la mitad de los textos y en la aplicación de unos determinados criterios de selección para la compilación de las muestras, estableciendo así unas bases para el diseño del corpus. Además, este proyecto fue precursor de un gran número de estudios sobre gramática inglesa basados en corpus.

La disponibilidad del ordenador a partir de 1960 originó un cambio radical en los estudios basados en corpus. El análisis por ordenador daba acceso a datos que previamente no era posible observar y que iban a aportar una perspectiva nueva sobre el lenguaje. El ordenador permitía almacenar enormes cantidades de textos y recuperar en una pantalla, una palabra, una cadena de palabras o fragmentos de textos con su contexto, rápidamente y de forma exhaustiva. El mismo *Survey of English Usage* Corpus se informatizó en cuanto fue posible.

A principios de los años ochenta, los corpus dedicados a la investigación lingüística no eran muy abundantes, y éstos únicamente se podían procesar mediante unidades de procesamiento centrales, los denominados *mainframe*. El desarrollo de ordenadores más potentes y la aparición del CD-ROM propiciaron la proliferación de los estudios basados en corpus, de modo que, hacia finales de los noventa, numerosos proyectos estaban en funcionamiento. Algunos proyectos de gran envergadura se llevaban a cabo con fines comerciales por editores de diccionarios, al mismo tiempo que se realizaban estudios a pequeña escala como parte de tesis doctorales y proyectos de investigación.

Desde los años 20 existe la tradición de contar las palabras de los textos con el fin de identificar las más frecuentes y por lo tanto, las candidatas a ser introducidas en los materiales de enseñanza de lenguas. Aunque el interés inicial se dirigía especialmente al léxico, los corpus informatizados y el software disponible en cada momento han permitido, paso a paso, ampliar el foco de atención hacia secuencias de palabras, colocaciones léxicas y hacia las características gramaticales y estilísticas de un autor o género en particular. Así, la actividad se va orientando a la identificación de los rasgos lingüísticos que caracterizan los estilos periodístico, literario y científico.

En el campo del inglés con fines específicos, los estudios realizados para identificar las características de los registros especializados se han basado en el análisis del lenguaje utilizado en textos auténticos.

La primera publicación sobre las características de la prosa científico-técnica (Barber, 1962) presentaba una investigación preliminar enfocada al estudio del vocabulario y de la estructura verbal y oracional. El autor introducía criterios cuantitativos y consideraba las variables de frecuencia y distribución. Barber, a la luz de los listados de palabras de West y Thorndike en la lengua general, pretendía aportar un listado de las palabras utilizadas generalmente en el lenguaje científico-técnico. El propio autor reconocía que la cantidad de textos analizados en su estudio (23.400 palabras aproximadamente), no era suficiente para obtener resultados generalizables y alentaba a continuar con la investigación:

To obtain certainty about a useful English vocabulary for the overseas science students, we must simply do more work: we need to take texts from all the main fields of science and technology, with representation of all the important sub-fields, and discover, on the lines suggested above, what words are common in one sub-field only, what are common in a whole main field, and what are common in all or several main fields; to compile the final list, a system of credits would be desirable, to enable frequency of occurrence to be balanced against range. This would enable us to draw up a general science vocabulary on a rational basis, and also to draw up separate word-lists for special fields.

De esta propuesta se puede inferir cómo yacen en estado latente los principios básicos que fundamentarán la lingüística del corpus, cuya relación con el inglés para fines específicos surge casi a la vez que el propio reconocimiento de la prosa científico-técnica como tal.

Skehan (1981) percibió el gran potencial de la aplicación de las herramientas del corpus en el campo del inglés con fines específicos. Él mismo escribió un programa para analizar textos de economía, con el fin de identificar las proporciones de los diferentes tipos de vocabulario (Aston, 1996). Sin embargo, casi cuarenta años después de la propuesta de Barber, Biber manifiesta que pocos estudios han seguido un enfoque basado en el corpus, es decir, *“using a large, principled collection of texts and combining quantitative and computational techniques with qualitative interpretations.”* (Biber, 1998).

Los estudios léxicos basados en corpus y los lenguajes especializados parecen converger en el estudio de la terminología, bajo el propósito de crear diccionarios especializados (Sager, 1990; Cabré, 1993; Pedersen, 1995; Meyer y Mackintosh, 1996; Pearson, 1998).

Respecto a la explotación de los corpus en la enseñanza de lenguas, existen dos modalidades: el profesor o investigador analiza el corpus para diseñar materiales y programas didácticos, o bien es el propio aprendiz quien manipula directamente el corpus. Este último procedimiento es impulsado por Tim Johns a principios de los 90, cuando manifiesta que *“the language-learner is also, essentially, a research worker whose learning needs to be driven by access to linguistic data”* (Johns, 1991). De aquí deriva el término *Data-driven Learning* (DDL) que denomina a este enfoque secundado, entre otros, por Aston, Ball, Barlow, Gavioli, etc.

McEnery y Wilson (1993) comienzan a testar la eficacia de los corpus manejados directamente por los estudiantes, obteniendo resultados satisfactorios en el aprendizaje. Unos años más tarde, estos autores consideran el uso de corpus en otras áreas más específicas, e identifican el IPA como una rama de la enseñanza de lenguas donde *“corpora can be used to provide many kinds of domain-specific material for language learning, including quantitative accounts of vocabulary and usage which address the specific needs of students in a particular domain more directly than those taken from more general language corpora”* (McEnery y Wilson, 1996).

Dudley-Evans y St. John (1998) admiten la utilidad de los corpus como fuente de recursos textuales para investigadores, profesores y estudiantes, especialmente los listados de palabras y concordancias, que facilitan examinar el contexto en el que aparece una palabra y seleccionar el vocabulario técnico y semi-técnico. Solamente cuestionan hasta qué punto un corpus es útil para el autoaprendizaje o para la enseñanza

de idiomas mediante el manejo directo del mismo en el aula. Aunque los autores no traten la cuestión en profundidad, su inclusión en este trabajo se considera relevante, puesto que asumen el análisis del lenguaje por ordenador como uno de los principales avances en el área del IPA en los noventa.

La Lingüística del Corpus se fundamenta en fuentes de evidencia textual para mejorar la descripción de la estructura y el uso del lenguaje humano, pero además, contribuye al desarrollo de diferentes campos, como la confección de herramientas lingüísticas informatizadas, la traducción automática y el reconocimiento y síntesis del habla. En estas áreas, las diversas aplicaciones de un corpus escrito dan lugar a diccionarios y correctores de textos informatizados. A su vez, los corpus orales desempeñan una función esencial para el entrenamiento y la validación de los sistemas de reconocimiento automático del habla, conversión texto a voz y codificadores de voz, cuyas aplicaciones comprenden desde servicios telefónicos automatizados, hasta ayudas para personas discapacitadas.

4.2 Definición y tipología de corpus lingüístico.

4.2.1 Definición de corpus lingüístico.

No todas las compilaciones o colecciones de textos se consideran corpus lingüísticos. Para pertenecer a esta categoría, el conjunto de textos tiene que cumplir unas determinadas características y los textos se deben seleccionar y ordenar siguiendo unos criterios previamente definidos. Las implicaciones que el concepto de corpus lingüístico conlleva, según se utiliza en la lingüística actual, se van haciendo patentes en el transcurso de su historia. Las primeras definiciones mencionan la investigación lingüística como finalidad que caracteriza a este cuerpo de textos, mientras que las siguientes van perfilando el concepto paso a paso.

En el proyecto Cobuild, Renouf se refiere al término corpus como: “*a collection of texts, of the written or spoken word, which is stored and processed on computer for the purposes of linguistic research*” (Sinclair, 1987).

EAGLES (1996), grupo experto de asesores sobre estándares en ingeniería del lenguaje, precisa que las muestras se deben seleccionar siguiendo unos criterios lingüísticos explícitos, y advierte algunas características informáticas para la codificación e identificación de los datos:

- a) *A corpus is a collection of pieces of language that are selected and ordered according to explicit linguistic criteria in order to be used as a sample of the language.*

- b) *A computer corpus is a corpus which is encoded in a standardised and homogenous way for open-ended retrieval tasks. Its constituent pieces of language are documented as to their origins and provenance.*

En una definición más reciente, Sinclair (2005) puntualiza la dimensión representativa que debe adquirir un corpus: “*A corpus is a collection of pieces of language text in electronic form, selected according to external criteria to represent, as far as possible, a language or language variety as a source of data for linguistic research.*”

Al unir y completar todas las definiciones consultadas (Johansson, 1991; Atkins, Clear y Ostler, 1992; McEnery y Wilson, 1996; Biber, Conrad y Reppen, 1998; etc), se podría concluir que un corpus lingüístico es una colección de textos, procedentes del lenguaje natural oral o escrito, seleccionados bajo unos determinados criterios, que caracterizan un estado, variedad o la totalidad del lenguaje, y que se utiliza para la investigación lingüística procesado por ordenador.

Con el fin de concretar el sentido que en este trabajo se utiliza el concepto de corpus, se hace constar una definición más detallada, que especifica las características generales del término:

Un corpus lingüístico es un conjunto de datos lingüísticos (pertenecientes al uso oral o escrito de la lengua, o a ambos), sistematizados según determinados criterios, suficientemente extensos en amplitud y profundidad de manera que sean representativos del total del uso lingüístico o de alguno de sus ámbitos, y dispuestos de tal modo que puedan ser procesados mediante ordenador con el fin de obtener resultados varios y útiles para la descripción y el análisis (Sánchez, 1995).

4.2.2 Tipos de corpus.

Partiendo del concepto de corpus según se ha explicado en el apartado anterior, se pueden encontrar diferentes tipos, dependiendo de la finalidad para la que son creados.

Un corpus general es un conjunto de textos analizado para contestar a determinadas preguntas planteadas sobre el vocabulario, gramática o estructura discursiva del lenguaje. Este corpus suele ser equilibrado, es decir, contiene textos de géneros y áreas diferentes, persiguiendo que se correspondan con el uso típico del lenguaje. De este modo, se puede utilizar como corpus de referencia en estudios comparativos.

Los corpus diseñados con un propósito específico o para determinados proyectos de investigación se denominan corpus especializados. Éstos se recopilan para fines tales como la creación de diccionarios, el estudio de la evolución del lenguaje de los niños, el estudio del lenguaje utilizado en un campo específico (medicina, ingeniería, economía, etc.), incluyendo los destinados al estudio de variaciones regionales o sociolingüísticas.

Dentro de los corpus bilingües o multilingües se encuentra el corpus paralelo y el corpus comparable. Un corpus paralelo contiene textos procedentes de una lengua y su traducción en otra. Por ejemplo, el *Canadian Hansard Corpus* está formado por las actas del parlamento canadiense en inglés y francés. Por otro lado, un corpus comparable consiste en dos o más corpus en lenguas distintas, cuyos textos se eligen para corresponderse uno con otro lo máximo posible. Permite comparar diferentes lenguas o variedades en circunstancias similares de comunicación, pero evitando las posibles distorsiones introducidas en la traducción del corpus paralelo.

El objetivo del denominado corpus de referencia es ser lo suficientemente grande como para representar todas las variedades del lenguaje y su vocabulario característico, de tal modo que pueda ser utilizado como base para la elaboración de diccionarios, tesauros, gramáticas y materiales de referencia fiables.

Sinclair (1992) define el corpus monitor o corpus dinámico como: “*huge, changing bodies of language of no finite size, flowing across a set of filters which extract linguistic evidence.*” Se trata de un corpus al que se van añadiendo constantemente materiales nuevos, a la vez que se eliminan cantidades equivalentes de material antiguo, permitiendo observar cambios recientes en el uso de la lengua. Con el progreso de los ordenadores y la posibilidad de almacenar grandes cantidades de datos, el corpus guarda el material antiguo a la vez que incorpora el nuevo, generándose así los corpus diacrónicos. De este modo, se pueden identificar palabras nuevas, el cambio en el uso de las palabras y cómo han evolucionado, nuevas acepciones, etc.

Los corpus también pueden ser clasificados como corpus no anotados o anotados. Los primeros se limitan a guardar el texto plano, es decir, el total de secuencias de palabras y signos de puntuación. En los corpus anotados, el ordenador relaciona con el texto plano cualquier tipo de información extra que se desee adherir.

4.3 Corpus principales.

La historia de los corpus lingüísticos se desarrolla en paralelo con la evolución de la informática. Los adelantos tecnológicos incorporados a los ordenadores han significado nuevos avances en las características de los corpus, en las herramientas de análisis y, por consiguiente, en los resultados obtenidos tras la investigación lingüística. En este sentido es posible distinguir diferentes generaciones de corpus dependiendo de dos variables: el incremento del volumen de datos y la anotación de los corpus.

Desde una perspectiva histórica existen tres generaciones de corpus. El transcurso de las generaciones viene marcado por el aumento de la potencia de los ordenadores que permiten aumentar el volumen de las muestras almacenadas. Sin embargo, el factor de

la anotación clasifica a los corpus en sólo dos generaciones: la primera generación se corresponde con los corpus no anotados y la segunda con los corpus anotados.

Un repaso cronológico de los proyectos de corpus más conocidos o más utilizados en estudios lingüísticos de la lengua inglesa puede aportar una idea de la diversidad y trascendencia alcanzada.

4.3.1 Corpus de primera generación.

La primera generación data de la década de los 60, cuando los corpus comienzan a ser procesados por ordenador y el volumen de compilación de datos apenas supera el millón de palabras. En esta generación son destacables los siguientes corpus:

The Brown Corpus: The Standard Sample of Present-Day American English.

En 1961 N. Francis y H. Kucera estuvieron a cargo de la elaboración del corpus compuesto por un millón de palabras (1.010.312). Las muestras consistían en fragmentos de 2.000 a 5.000 palabras de textos en prosa, de distintos estilos y variedades, editados en los Estados Unidos durante 1961. Los autores de los textos eran hablantes nativos de inglés americano.

El uso de la palabra *standard* en el título del corpus no significa que se considere como muestra del inglés americano estándar, sino como corpus de referencia para estudios comparativos que necesiten la misma fuente de datos.

En la actualidad existe una serie de corpus diseñados bajo los mismos criterios pero que recopilan una variedad o lengua diferente, denominados *Brown Family*: Frown (Inglés americano, 1991-92), LOB (Inglés británico, 1961), Pre-LOB (Inglés británico, 1931+/- 3 años), FLOB (Inglés británico, 1991-92), Kolhapur (Inglés de la India, 1978), ACE (Inglés australiano, 1986), WWC (Inglés de Nueva Zelanda, 1986-1990) y LCMC (Chino mandarín, 1991+/- 3 años).

The London-Lund Corpus of Spoken English (LLC).

Este corpus deriva de dos proyectos: *The Survey of English Usage (SEU)* propuesto por R. Quirk en la Universidad de Londres (1959) y *The Survey of Spoken English (SSE)*, propuesto por J. Svartvik en la Universidad de Lund (1975).

El objetivo de Quirk fue elaborar un corpus del inglés británico hablado y escrito para describir la gramática de los hablantes adultos cultos. Sin embargo, el hecho de que un 50% fueran datos lingüísticos orales supuso una gran dificultad. Más tarde Svartvik consiguió completar y digitalizar esta parte oral hasta alcanzar un millón de palabras. Además, el corpus incluía anotaciones derivadas del análisis prosódico del texto oral. Esto propició especialmente la realización de numerosos estudios sobre la estructura y funciones del discurso.

The Lancaster-Oslo-Bergen Corpus of British English (LOB).

Tres instituciones toman parte en la creación del LOB: la Universidad de Lancaster, la Universidad de Oslo y el Centro de Informática Noruego para Humanidades de Bergen; durante 1970-1976 el proyecto fue dirigido por G. Leech en la Universidad de Lancaster y durante 1977-1978 por S. Johansson en la Universidad de Oslo.

El propósito del proyecto consistía en elaborar un corpus del inglés británico equivalente al *Brown Corpus*, con el fin de obtener una representación general de tipos de textos con fines de investigación lingüística. Para facilitar la combinación del uso de ambos corpus, intentaron compaginar en lo máximo los materiales del corpus británico con los del americano.

En esta etapa de la lingüística del corpus, sería conveniente hacer notar que los compiladores del *Lob Corpus*, reconocieron como insuficiente la cantidad de un millón de palabras para representar adecuadamente la incidencia de las unidades o fenómenos lingüísticos de baja frecuencia.

4.3.2 Corpus de segunda generación.

La segunda generación de corpus se puede situar a mediados de la década de los 80, cuando el avance tecnológico introduce un nuevo elemento: el reconocedor óptico de caracteres (OCR). Éste permite digitalizar los textos y agilizar el proceso de captación de muestras. Por consiguiente, el aumento del volumen de datos recopilados es considerable. Además, en esta etapa se consigue un logro muy importante: la expansión de los corpus como herramientas de estudio y análisis.

The Longman-Lancaster English Language Corpus.

Este corpus es el resultado del trabajo en colaboración entre *Longman Publishers* y la Universidad de Lancaster. El corpus consta de unos 30 millones de palabras del inglés escrito, procedentes de textos publicados.

Una parte de las muestras se extrajo de obras literarias famosas y otra parte de libros de ficción menos conocidos, seleccionados al azar. Respecto a las muestras que no pertenecen al género de ficción, éstas se recopilaron de varias fuentes, como las ciencias sociales y naturales, el comercio y las finanzas, las artes, el ocio, etc.

El objetivo de este proyecto era ofrecer una serie de publicaciones fundamentadas en los resultados del análisis del corpus. De esta manera, los resultados serían accesibles al público no especializado, por medio de diccionarios, libros de gramática y cursos de inglés para extranjeros.

Proyecto Cobuild.

A principios de los años 80, John Sinclair dirige el proyecto Cobuild en la Universidad de Birmingham con la colaboración de la editorial Collins. La finalidad del proyecto consistía en recopilar un corpus de textos almacenados en ordenador (7 millones de palabras), para la elaboración de un diccionario y el estudio de la lengua. El principal resultado consistió en la producción del diccionario de la lengua inglesa “Collins-Cobuild”, al cual sucedieron numerosas publicaciones dedicadas a la enseñanza, aprendizaje y estudio de la lengua inglesa.

Posteriormente la editorial decidió aumentar el tamaño del corpus hasta conseguir los 200 millones de palabras, convirtiéndose en *The Bank of English Corpus*, que pertenece a la tercera generación.

4.3.3 Corpus de 3ª generación.

A finales de la década de los 90 se emprenden proyectos de gran envergadura que dan lugar a los “Mega-corpus” o corpus de tercera generación. Ésta se caracteriza por las ingentes cantidades de datos compilados que el avance de la tecnología permite almacenar, procesar y analizar de forma cada vez más rápida y exhaustiva. La mayoría de estos proyectos están respaldados por importantes editoriales y son destinados igualmente a fines comerciales.

The Bank of English Corpus.

The Bank of English tiene su origen en el proyecto Cobuild, cuyo corpus se amplía hasta alcanzar un volumen de 200 millones de palabras, y además es anotado morfológica y sintácticamente. El corpus continúa creciendo con la constante adición de material nuevo, de modo que en el año 2005 cuenta con 525 millones de palabras.

La mayoría de las muestras datan de 1990 en adelante y provienen tanto del inglés escrito como del hablado, cubriendo una amplia variedad de modalidades y fuentes. Los textos escritos proceden de periódicos, revistas, libros de ficción, publicidad, cartas, informes, etc. La parte oral procede de transcripciones de conversaciones espontáneas de la vida diaria, entrevistas, debates, transmisiones de radio, etc.

Todas las muestras se almacenan en una base de datos accesible a los investigadores. Los lexicógrafos y lingüistas implicados en este proyecto buscan patrones de combinaciones de palabras, comprueban la frecuencia de las palabras, identifican todos los usos de una determinada palabra y analizan los resultados para que la información registrada en los diccionarios y trabajos de referencia sea fiable y esté respaldada por los hechos. El corpus también es de gran utilidad para profesores de inglés, traductores y estudiantes, pues supone una gran fuente de recursos para sus estudios y actividad profesional.

Las investigaciones llevadas a cabo por Cobuild en los últimos veinte años han demostrado que tales compilaciones de textos son necesarias para realizar estudios lingüísticos fiables y válidos.

The British National Corpus (BNC).

Oxford University Press dirige el desarrollo de este proyecto (1991-1994) en colaboración con las editoriales Longman y Larousse, las Universidades de Lancaster y Oxford, la *British Academy* y la *British Library*.

The British National Corpus consta de 100 millones de palabras del inglés moderno oral (10 %) y escrito (90 %). La parte oral está compuesta tanto por muestras de la lengua hablada recogida en toda clase de contextos diferentes, como por una gran cantidad de conversaciones informales auténticas. Estas conversaciones son producidas de forma espontánea, sin guión, por voluntarios previamente seleccionados, siguiendo un criterio de variedad regional, social y edad. La parte escrita incluye una gran variedad de tipos de texto y variedades del lenguaje, como extractos de periódicos regionales y nacionales, libros académicos y de ficción, cartas y memoranda, etc.

The International Corpus of English (ICE).

El principal propósito de la elaboración de este corpus es recopilar el material para realizar estudios comparativos de las variedades internacionales del inglés. El proyecto fue iniciado en 1990 por la Universidad de Londres, a la cual se unieron un total de veinte centros de investigación, con el fin de construir un corpus electrónico de su propia variedad nacional o regional del inglés.

Todos los corpus siguen el mismo diseño y esquema para su anotación gramatical. Cada corpus consta de un millón de palabras pertenecientes al uso oral y escrito. Todas las muestras son posteriores a 1989, cuyos autores o hablantes son hombres y mujeres adultos, nativos o que recibieron su educación en la variedad del inglés correspondiente.

Los corpus están siendo anotados a diferentes niveles para aumentar su valor en la investigación lingüística: marcas de texto (*Textual Markup*), etiquetado por clases de palabras (*Wordclass Tagging*) y a nivel sintáctico (*Syntactic Parsing*).

The American National Corpus (ANC).

El proyecto *American National Corpus* persigue la compilación de un corpus representativo del inglés americano, producido de forma oral y escrita desde 1990 en adelante. En el año 2003 el volumen de datos recopilados alcanzó los 11 millones de palabras. No obstante, el objetivo establecido consiste en enriquecer el corpus hasta reunir 100 millones de palabras, siguiendo el modelo del *British National Corpus*, y conseguir un corpus comparable.

Después de alcanzar tal cantidad, se pretende seguir añadiendo muestras regularmente, del orden de un 10% de material nuevo cada cinco años y conferirle el carácter de corpus monitor. El ANC constituirá una fuente de recursos para la investigación lingüística y el desarrollo tecnológico.

Este proyecto es liderado por Randi Reppen con el apoyo de varias instituciones americanas como el ANC Consortium, el *TalkBank project* y *National Science Foundation*.

El siguiente cuadro ofrece una visión más amplia de corpus en lengua inglesa (aunque no aparecen todos), que pretende complementar la muestra de corpus seleccionados para ilustrar las tres generaciones:

<p>CORPUS MONITOR The global English Monitor Corpus</p>	<p>CORPUS DEL INGLÉS ACADÉMICO Y PROFESIONAL The Michigan Corpus of Academic Spoken English The British Academic Spoken English corpus The Reading Academic Text corpus The Academic Corpus The Corpus of Professional Spoken American English</p>
<p>CORPUS SINCRÓNICO The Longman Written American Corpus</p>	<p>CORPUS ORALES SEC, MARSEC and Aix-MARSEC The Oral Vocabulary of the Australian Worker Corpus The Bergen Corpus of London Teenage Language The Cambridge and Nottingham Corpus of Discourse in English The Spoken Corpus of the Survey of English Dialects The Intonational Variation in English Corpus The Longman British Spoken Corpus The Longman Spoken American Corpus The Santa Barbara Corpus of Spoken American English The Saarbrücken Corpus of Spoken English The Switchboard Corpus The Wellington Corpus of Spoken New Zealand English The Limerick corpus of Irish English The Hong Kong Corpus of Conversational English</p>
<p>CORPUS DIACRÓNICO The Helsinki Corpus of English Texts The ARCHER corpus The Lampeter Corpus of Early Modern English Tracts The Dictionary of Old English Corpus in Electronic Form Early English Books Online The Corpus of Early English Correspondence The Zurich English Newspaper Corpus The Innsbruck Computer Archive of Machine-Readable English Texts The Corpus of English Dialogues A Corpus of Late Eighteenth-Century Prose A Corpus of Late Modern English Prose The Cambridge International Corpus</p>	<p>CORPUS DE APRENDICES (Learner corpus) The Child Language Data Exchange System The Louvain Corpus of Native English Essays The Polytechnic of Wales corpus The International Corpus of Learner English The LINDSEI corpus The Longman Learners' Corpus The Cambridge Learner Corpus</p>
<p>CORPUS ANOTADOS The Lancaster Parsed Corpus The SUSANNE corpus The CHRISTINE corpus The LUCY corpus ICE-GB The Penn Treebank Parsed historical corpora</p>	

Cuadro 4.1 Otros corpus en lengua inglesa.

4.4 Corpus especializados.

La profusión de corpus generales no es comparable a la cantidad de corpus especializados y menos aún, a los corpus dedicados exclusivamente al inglés de las telecomunicaciones o relacionados con él.

El primer corpus del inglés científico-técnico, según hace constar Swales (2004), fue el recopilado por Huddleston para su estudio *The Sentence in Written English: A Syntactic Study Based on an analysis of Scientific Texts*, 1971. Este trabajo deriva de un proyecto financiado por el gobierno británico para investigar las características lingüísticas del inglés científico, que fue llevado a cabo por Huddleston, Hudson y Winter a mediados de los años 60. El informe final *Sentence and Clause in Scientific English* no está localizable, pero en el trabajo de Huddleston se puede obtener información sobre el corpus que recopilaron. Éste contenía 135.000 palabras procedentes de libros de texto, revistas especializadas y de divulgación, del ámbito de la física, química y biología.

Huddleston centra su investigación en la gramática común a las tres materias y aporta una descripción gramatical selectiva del corpus. Asimismo, alberga la esperanza de que su estudio sea útil para preparar cursos de inglés científico y deja la puerta abierta para continuar su trabajo.

Until further comparative work... is done one cannot of course tell how far the statistical properties of the corpus reported in the present work are peculiarly characteristic of written scientific English and how far they are generalizable to other varieties; I hope, however, to have provided a solid basis for such comparative study (Huddleston en Swales, 2004).

En la actualidad, la accesibilidad a corpus generales de tercera generación permite seleccionar la sección dedicada a la materia requerida, obteniendo una cantidad de datos suficiente para formar un subcorpus específico. Tal sería el caso del apartado del inglés científico-técnico en general. Sin embargo, cuando se delimita una rama de la ciencia o tecnología y el dominio es cada vez más limitado y específico, se hace necesario elaborar un corpus especializado con las muestras adecuadas, si se pretende que los resultados del estudio no sean meramente ilustrativos.

Numerosos centros han impulsado la investigación basada en corpus del inglés científico-técnico, en ocasiones por iniciativa académica, y en otras, como producto de la colaboración entre varias universidades, guiados especialmente por el interés de los profesores de lenguas en universidades donde se imparten estudios técnicos. Dentro de este contexto, se consideran los siguientes corpus específicos del inglés científico-técnico, relacionados parcialmente con el inglés de las telecomunicaciones, en el sentido de que comparten materias afines.

Entre los primeros corpus de inglés para fines académicos, pioneros en la categoría de corpus especializados y compuestos por textos científico-técnicos se encuentra el *Jiaotong Daxue English of Science and Technology Corpus (JDEST)*. El Instituto de Procesamiento de Lenguaje Natural de la Universidad de Shanghai Jiaotong promueve este proyecto en los años 80, bajo la dirección de Yang. Inicialmente el corpus constaba

de 1 millón de palabras, pero fue ampliado hasta conseguir un total de 4,5 millones de palabras.

Tomando los libros de texto como fuente de procedencia de las muestras coinciden el corpus HKUST (1992) y el SEEC (1999/2003). El primero, proyectado por Fang en el Centro de Lenguas de la Universidad de Ciencia y Tecnología de Hong Kong, reúne un millón de palabras del inglés utilizado en informática. El segundo contiene muestras de la ingeniería mecánica, termodinámica, mecánica e ingeniería de los materiales, procesos de manufacturación, dibujo técnico y programación de computadoras, materias obligatorias en la Universidad de Walailak (Tailandia). Moudraia desarrolla este proyecto con el fin de aportar unos fundamentos lo suficientemente sólidos como para elaborar un programa fiable basado en el léxico común. Este mismo objetivo promueve la investigación de Curado (2000), que recopila un corpus con muestras relacionadas con los estudios de Informática Técnica, Ingeniería Informática, Sonido e Imagen, Biblioteconomía y Documentación, y Comunicación Audiovisual. Su aportación estriba en la incorporación de textos procedentes tanto de fuentes académicas como del mundo profesional.

En esta misma línea, pero a mayor escala, el grupo de Tecnologías del Lenguaje de la Universidad Politécnica de Valencia diseña y recopila el corpus de textos de inglés científico-técnico y académico denominado ACIA. En la primera fase del proyecto el corpus alcanza tres millones de palabras (1999), aunque se continúan incorporando muestras con la intención de conseguir 20 millones de palabras. Los textos se clasifican según las disciplinas a las que pertenecen y el tipo de género (artículos de revistas científicas de investigación, informes técnicos, abstracts, tesis y comunicaciones presentadas en congresos). Las disciplinas consideradas son aquellas que han estimado como las más representativas de la UPV: agricultura, física, biología, química, electrónica, ingeniería civil, ingeniería industrial, ingeniería mecánica, informática y telecomunicaciones.

El único corpus del que se tiene constancia en esta investigación, que está relacionado exclusivamente con el lenguaje utilizado en las telecomunicaciones es el corpus CRATER: *Corpus Resources and Terminology Extraction*. Este corpus es producto de la cooperación entre la Universidad de Lancaster y la Universidad Autónoma de Madrid para un proyecto europeo, donde se encargan de recoger los textos originarios de la *International Telecommunications Union* en lengua inglesa, francesa y española. Aunque este corpus cubre el lenguaje producido en el ámbito profesional de las telecomunicaciones, carece del lenguaje utilizado como medio de enseñanza, es decir, no contiene textos procedentes de libros de texto.

Como conclusión se puede afirmar que, en el transcurso de este trabajo, no se ha hallado ningún corpus disponible que satisfaga los requisitos necesarios para realizar el estudio aquí propuesto. No ha sido posible localizar un corpus exclusivo y representativo del lenguaje relacionado con los estudios de Ingeniería de Telecomunicaciones y Telemática y su correspondiente ejercicio profesional.

4.5 Consideraciones relevantes para la creación de un corpus.

Los estudios basados en corpus lingüísticos parten de la selección del corpus que se va a utilizar. Si no existe uno disponible y afín a los propósitos de la investigación, la primera fase debe centrarse en el diseño y la elaboración de un corpus adecuado, compilado con una finalidad en particular.

Desde los años sesenta, compilación, estructura y tamaño de los corpus han sido tema de continua atención entre los lingüistas del corpus, ya que estos factores afectan directamente a la validez y fiabilidad de la investigación. Diferentes autores ofrecen sus recomendaciones y directrices, consideradas cruciales, y de las cuales afirman que dependerá el grado de éxito y fiabilidad del estudio. Sinclair (1991) aporta unos criterios que guían la creación de un corpus general, en torno al contenido del corpus, la selección y organización de las muestras, y que son igualmente aplicables para un corpus específico.

De acuerdo con Pearson (1998), en la bibliografía disponible relacionada con los corpus de fines específicos, los autores se limitan a informar sobre el proceso de compilación y no argumentan en profundidad las pautas seguidas para el diseño. Por consiguiente, la autora combina los principios utilizados con el fin de obtener unos criterios base para la composición de un corpus específico.

Posteriormente, Sinclair propone y define más detalladamente diez principios a seguir en el diseño de un corpus general y la recopilación de muestras (Wynne, 2005). El decálogo es presentado a continuación y será considerado en los apartados siguientes.

1. *The contents of a corpus should be selected without regard for the language they contain, but according to their communicative function in the community in which they arise.*
2. *Corpus builders should strive to make their corpus as representative as possible of the language from which it is chosen.*
3. *Only those components of corpora which have been designed to be independently contrastive should be contrasted.*

4. *Criteria for determining the structure of a corpus should be small in number, clearly separate from each other, and efficient as a group in delineating a corpus that is representative of the language or variety under examination.*
5. *Any information about a text other than the alphanumeric string of its words and punctuation should be stored separately from the plain text and merged when required in applications.*
6. *Samples of language for a corpus should wherever possible consist of entire documents or transcriptions of complete speech events, or should get as close to this target as possible. This means that samples will differ substantially in size.*
7. *The design and composition of a corpus should be documented fully with information about the contents and arguments in justification of the decisions taken.*
8. *The corpus builder should retain, as target notions, representativeness and balance. While these are not precisely definable and attainable goals, they must be used to guide the design of a corpus and the selection of its components.*
9. *Any control of subject matter in a corpus should be imposed by the use of external, and not internal, criteria.*
10. *A corpus should aim for homogeneity in its components while maintaining adequate coverage, and rogue texts should be avoided.*

4.5.1 Representatividad.

El aspecto principal en el que coinciden todos los autores es la representatividad (Renouf et al., 1987; Biber et al., 1993; Sánchez, 1995; Sánchez y Cantos, 1997; EAGLES, 1996; McEnery y Wilson, 1996; Pearson, 1998; Sinclair, 1991 y 2005; etc). Pretender que un conjunto de muestras del lenguaje represente el uso total de una variedad de la lengua es una cuestión bastante problemática. La controversia procede de la misma complejidad que supone definir la representatividad, y conseguir que una fracción contenga los componentes que le confieran el carácter representativo de un total. Biber define representatividad como “*the extent to which a sample includes the full range of variability in a population*” (Biber 1993).

Por medio del corpus se acota un lenguaje específico integrado por una variedad de constituyentes que deberán ser identificados y captados. Por lo tanto, el diseño del corpus deberá ser extremadamente cuidadoso y los criterios fijados fielmente seguidos, con el fin de que los datos recopilados reflejen las características del objeto de estudio.

A pesar de los intensos esfuerzos que se realicen, la representatividad de un corpus sólo llegará á ser aproximada, pues no es posible que una compilación limitada de textos represente completamente el lenguaje de toda una comunidad de hablantes.

Leech (1991) sugiere que un corpus podría considerarse representativo cuando los hallazgos obtenidos de su análisis se puedan hacer generalizables al resto del lenguaje que tipifica.

4.5.2 Tamaño.

El tamaño del corpus está estrechamente relacionado con la representatividad. Aunque no se ha determinado qué cantidad de palabras son necesarias para que un corpus sea representativo, existen diferentes opciones.

Según Pearson (1998), un millón de palabras es el tamaño normalmente seleccionado para un corpus específico. Las razones oscilan entre, simplemente la intuición, hasta conclusiones procedentes de proyectos de mayor envergadura, donde resuelven que un millón de palabras representa una proporción razonablemente alta de la parte del lenguaje bajo estudio.

Por otro lado, Sinclair (1991) señala que un corpus de 10-20 millones de palabras puede constituir un pequeño corpus útil, pero que no será adecuado para una descripción fiable de la totalidad del lenguaje. Por muy grande que sea un corpus, nunca será más que una muestra minúscula de todo el lenguaje utilizado por todos los hablantes de una lengua. Sinclair aconseja que un corpus debería ser lo más grande posible y seguir creciendo, para poder estudiar el comportamiento de las palabras en los textos, ya que éstas se distribuyen de forma irregular y la mayoría aparecen sólo una vez.

Kennedy (1998) sugiere calcular cuántas ocurrencias de una palabra son necesarias para una descripción idónea. En general, del 40 al 50% de los lemas en un corpus ocurren sólo una vez, y se necesita más de una ocurrencia para establecer una base de descripción. Por otro lado, una palabra de alta frecuencia como la preposición *at*, normalmente aparece unas 5.500 veces en un millón de palabras, cantidad más que suficiente para fines descriptivos.

Aunque para conseguir una descripción idónea de fenómenos de baja frecuencia como las colocaciones, se necesitan corpus de gran tamaño, no tendría sentido agrandar cada vez más el corpus si no se puede trabajar con los datos. Según Kennedy (1998) un corpus de gran tamaño no necesariamente representa un lenguaje o una variedad mejor que otro más pequeño, y afirma: “*At this stage we simple do not know how big a corpus needs to be for general or particular purposes*”.

Pearson (1998) añade que no hay ninguna justificación para poner límites al tamaño de un corpus. El límite viene dado por la cantidad de textos disponibles o convertibles a formato electrónico y que a su vez satisfacen los criterios de selección adoptados.

4.5.3 Variedad.

Análisis basados en corpus demuestran que existen diferencias importantes en el uso de características léxicas, gramaticales y discursivas entre diferentes variedades del lenguaje. Por eso es necesario introducir variedad de registros, de autores, variedades geográficas, variedad temática, etc., extrayendo las muestras del lenguaje de diferentes fuentes.

4.5.4 Cronología.

El criterio temporal define el intervalo de tiempo en el que se han producido las muestras, es decir, el periodo de tiempo que cubre el corpus. En este sentido un corpus puede ser sincrónico o diacrónico.

Un corpus sincrónico es una colección estática de textos, seleccionados según unos principios y recopilados con la intención de que sean característicos del lenguaje o un aspecto del lenguaje, en un periodo de tiempo en particular.

Un corpus diacrónico debe cubrir de forma sistemática diferentes periodos de tiempo, para asegurar los datos necesarios en los que basar estudios de cambios en la lengua.

En el caso de un corpus específico utilizado para estudios terminológicos es conveniente que las muestras se hayan producido en los últimos diez años (Pearson, 1998).

4.5.5 Tipos de texto o géneros.

Un procedimiento clave para conseguir representar la variedad lingüística es introducir muestras de todos los tipos de texto o géneros del lenguaje. Para identificar los textos característicos de una comunidad de hablantes se deben tomar como referencia las situaciones comunicativas que acontecen entre ellos.

Finalmente, es importante ser realista ya que todos los corpus tienen sus limitaciones, bien temporales, informáticas, financieras, de disponibilidad de textos, etc. Sin embargo, un corpus bien diseñado ofrece la oportunidad y es útil para investigar la variedad o aspectos lingüísticos propuestos.

4.5.6 Anotación del corpus.

La anotación o etiquetado de un corpus consiste en añadir información lingüística de diversa índole en forma de etiquetas adheridas al texto. Existen numerosos tipos de anotación: fonética, semántica, pragmática, discursiva, estilística, etc. La forma de anotación más frecuente es la anotación morfológica, que aporta información sobre las palabras y su morfología. Tal etiquetado es fundamental para poder hacer búsquedas más precisas en el corpus, y es además, el estadio previo a otros tipos de anotación más sofisticado, como el análisis sintáctico o *parsing*.

Toda clase de información que pueda resultar útil para los propósitos de la investigación se puede añadir mediante la anotación. Este procedimiento se realiza de forma manual o automáticamente con la ayuda del software adecuado. En cualquier caso, se debe utilizar un sistema de codificación estándar que permita el transporte y reutilización de los textos.

El etiquetado lingüístico es una aplicación opcional que proporciona un valor añadido al corpus. Sin embargo, la codificación de las muestras con información meta-textual o metadatos, se convierte en un requisito indispensable. Los metadatos son información descriptiva y altamente estructurada sobre el contenido y las características de los textos almacenados. Estos datos son cifrados por medio del código que encabeza cada texto, y que facilita tanto su procesamiento como su análisis. El código permite recuperar la información suficiente sobre una muestra para poder localizar la fuente textual completa de donde deriva. Dependiendo del software que haya procesado el corpus, el código se hace visible en pantalla, tal y como se explica más adelante.

CAPÍTULO 5

DISEÑO Y DISTRIBUCIÓN DEL CORPUS DE TELECOMUNICACIONES

5.1 Diseño del Corpus.

El corpus específico recopilado para servir a los propósitos de esta investigación, se ha diseñado cuidadosamente con la intención de ser considerado como razonablemente representativo del uso escrito que los ingenieros de telecomunicaciones hacen de la lengua. Para ello, se ha intentado seguir fielmente, y en la medida de lo posible, tanto los criterios propuestos para la selección de textos, como las indicaciones recomendadas para la recopilación de un corpus general. Estas directrices se han transferido al ámbito específico y se han complementado con las indicaciones de los especialistas.

De acuerdo con los principios propuestos por Sinclair (Wynne, 2005), representatividad y equilibrio deben constituir un objetivo primordial aunque no se puedan alcanzar ni definir con precisión. No obstante, estas nociones se mantendrán como guía en el diseño del corpus y en la selección de sus componentes.

A fin de que los resultados del análisis e investigación sean fiables y extrapolables, las muestras recopiladas proceden de actos de comunicación reales, se han preparado para su procesamiento por ordenador y se han sistematizado en relación con los siguientes criterios: variedad temática, cronología, geografía, modalidad y tamaño.

Teniendo en cuenta todos los criterios anteriores, se estima que el corpus de telecomunicaciones recopilado para esta investigación supone una aportación considerable al estudio de este lenguaje de especialidad.

5.1.1 Variedad temática.

Como guía en la búsqueda de una representatividad temática dentro del amplísimo campo de las telecomunicaciones, y puesto que esencialmente este trabajo está destinado al enriquecimiento y beneficio de los estudiantes de ingeniería, se han tomado como referencia los planes de estudios de Ingeniería de Telecomunicaciones y de Telemática en la Universidad Politécnica de Cartagena.

Es necesario puntualizar que cada área de conocimiento que integra los planes de estudio, presenta una unidad temática de referencia para buscar muestras. Sin embargo,

en ningún caso existe una predisposición para recopilar sólo muestras del lenguaje académico. La intención es recoger, de manera aleatoria, textos procedentes tanto del ámbito académico como del profesional.

5.1.2 Cronología.

El criterio temporal define el intervalo de tiempo en el que se han producido las muestras. El corpus de telecomunicaciones es un corpus sincrónico que contiene muestras que datan principalmente del periodo de tiempo comprendido entre los años 1997 y 2005. El proceso de recopilación de textos se ha realizado en dos fases.

Fase I.

La primera fase comienza en el año 2000. En esta fecha se recogen los textos encontrados al tomar como referencia temática las áreas de conocimiento de los tres primeros años de los estudios superiores, y los tres años de los estudios técnicos. Las muestras recopiladas en esta fase se han producido principalmente entre 1997 y 2003.

Fase II.

La segunda fase se inicia en el 2004. Los textos introducidos están relacionados con los dos últimos años de los estudios superiores, es decir, caracterizan las dos intensificaciones o especialidades disponibles.

En esta etapa se procede igualmente a completar y equilibrar el resultado de la primera fase. Desde el comienzo de la recopilación se han ido agregando muestras hasta alcanzar un tamaño considerable, momento en el cual se detiene la compilación para realizar el análisis. No obstante, la intención es conferirle un carácter dinámico al corpus y seguir recopilando muestras para que no se convierta en obsoleto, puesto que la información extraída de un corpus específico, de un período de tiempo en concreto, es válida mientras la base textual se mantenga en uso.

La norma de selección es captar siempre los textos elaborados más recientemente. Sin embargo, si existe un texto considerado básico para una materia y data de años anteriores, es igualmente introducido en el corpus debido a su relevancia para la formación del ingeniero.

5.1.3 Ámbito Geográfico.

Respecto al ámbito geográfico, al tratarse de un corpus de la lengua inglesa, se recogen principalmente muestras del inglés británico y americano. El inglés de Canadá, Australia y otras variedades geográficas no está contemplado en el corpus, aunque al comienzo del diseño se tomaron en consideración. Desafortunadamente se desechó su compilación debido a la escasez de muestras disponibles por tratarse de variaciones

geográficas minoritarias. Existe un número mucho mayor de publicaciones y páginas web de las variedades británica y norteamericana, especialmente de ésta última, que del resto de variedades. Las publicaciones norteamericanas suelen tener una mayor difusión y ejercen una influencia dominante, uno de los factores que les ha permitido establecerse como destacados representantes de las telecomunicaciones.

Una tercera división comprende el inglés producido por los hablantes no nativos, considerado en el corpus para su posterior análisis y contraste, puesto que una gran producción científica y tecnológica es expresada en inglés por hablantes no nativos para su difusión mundial.

La mayoría de los documentos recopilados aportan información sobre su procedencia, de lo contrario, el origen se determina por la ubicación de la página web en la que aparecen o el país donde son publicados.

5.1.4 Modalidad.

El corpus refleja solamente la modalidad escrita del uso lingüístico. Resulta muy complicado reunir muestras orales auténticas del ámbito de las telecomunicaciones, y las halladas son tan escasas que el resultado de su análisis no sería significativo. El material disponible está compuesto por diálogos preparados para ejercicios de comprensión oral y transcripciones de noticias tecnológicas. En ambos casos se sigue un guión y por lo tanto, las muestras no se corresponden con la producción espontánea de la comunicación oral entre personas.

Los textos escritos recopilados son documentos completos, es decir, no son fragmentos de textos más largos sino que el original se introduce en su totalidad. Por consiguiente, las muestras no comparten el mismo tamaño. Los libros se recogen por capítulos o secciones completas. En la sección de *abstracts* se guardan los resúmenes individuales, es decir, aquellos que aparecen por separado y no como parte de artículos de investigación. Éstos se almacenan por completo, incluyendo sus respectivos resúmenes y exceptuando las expresiones gráficas y la sección de bibliografía.

Sin embargo, hay ocasiones en las que es complicado delimitar el texto, como es el caso de las páginas web. Para acotar el texto se han seguido dos opciones: bien se incluye el texto que aparece en pantalla, bien se selecciona siguiendo las diferentes secciones hacia donde dirige el enlace asociado.

5.1.5 Tamaño.

El corpus de Telecomunicaciones contiene 5,5 millones de palabras (5.533.705), es decir, de secuencias de caracteres separadas por espacios en blanco. Las muestras relacionadas con los tres primeros años de los estudios superiores y los tres años de los

estudios técnicos alcanzan un volumen de 3.652.548 de palabras. Las muestras relacionadas con las especialidades añaden 1.881.157 de palabras.

El resultado obtenido está condicionado por factores heterogéneos. Fundamentalmente, se ha intentado recopilar una cantidad similar de muestras para cada sección temática y medio de difusión, con el fin de obtener un corpus equilibrado. No obstante, en ocasiones no ha sido posible hallar el mismo volumen para cada apartado, debido a la diversa naturaleza de la propia disciplina y las subdisciplinas que la conforman, a la disponibilidad de textos en formato electrónico, a que no toda la información se difunde por los mismos medios ni en proporciones exactas, etc.

El volumen total conseguido, aunque no alcance los cientos de millones de palabras, quintuplica el tamaño sugerido para un corpus especializado en las primeras propuestas – 1 millón de palabras (Kennedy, 1998; Pearson, 1998; Curado, 2001). Cuantos más datos se recojan, más clara y precisa será la imagen que se obtendrá del lenguaje. Para los objetivos de esta investigación, se considera que cinco millones de palabras constituyen un tamaño aceptable, en relación con la especificidad temática y la variedad de textos considerados. La única condición sine qua non requerida respecto al volumen del corpus es que debe alcanzar un tamaño muestral adecuado. En este sentido, cinco millones de palabras aseguran niveles de confianza en los resultados de inferencia superiores al 99%.

5.2 Distribución del Corpus.

La estructura física del corpus, previa a su procesamiento, está jerárquicamente organizada, de modo que en una carpeta principal denominada “Corpus de Telecomunicaciones”, se incluyen tres grandes archivos que comprenden las muestras recogidas de cada área geográfica:

1. Inglés americano (*USA English*: US)
2. Inglés británico (*British English*: BE)
3. Inglés de los no nativos (*Non-native English*: NN)

Variedad geográfica	Número de palabras
Inglés americano	3.513.282
Inglés británico	874.624
Inglés de hablantes no nativos	1.145.799
Total	5.533.705

Cuadro 5.1 Volumen del corpus distribuido por variedades geográficas

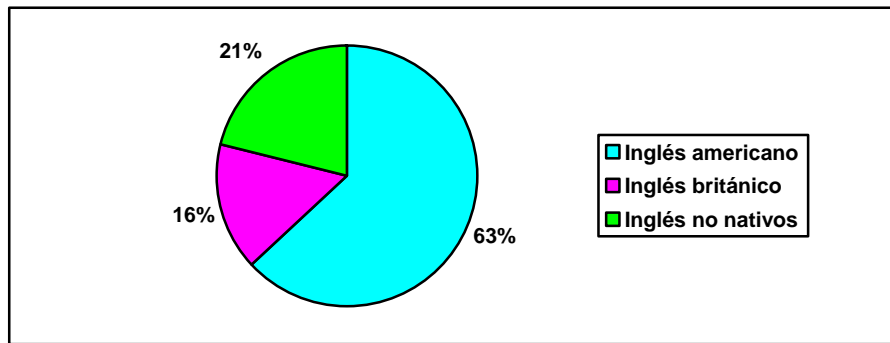


Figura 5.1 Volumen del corpus distribuido por variedades geográficas.

Cada archivo se subdivide, a su vez, en las diferentes áreas de conocimiento contempladas en los planes de estudios conducentes a la obtención de las titulaciones de Ingeniero Superior de Telecomunicación y de Ingeniero Técnico de Telecomunicación, Especialidad en Telemática. Cada área se organiza en diferentes carpetas con las asignaturas correspondientes a cada una de dichas áreas.

1. Electrónica (*Electronics*: 01):
 - Componentes Electrónicos (*Electronic components*: 01)
 - Electrónica Analógica (*Analogue Electronics*: 02)
 - Electrónica Digital (*Digital Electronics*: 03)
 - Fotónica (*Photonics*: 04)
2. Arquitectura y Tecnología de Computadoras (*Computing Architecture and Technology*: 02):
 - Fundamentos de Computadores (*Computing Fundamentals*: 01)
 - Sistemas Electrónicos Digitales (*Digital Electronic Systems*: 02)
3. Ingeniería Telemática (*Telematic Engineering*: 03):
 - Proyectos (*Projects*: 01)
 - Telemática (*Telematics*: 02)
 - Sistemas de Información Distribuidos (*Distributed Information Systems*: 03)
 - Software de Comunicaciones (*Communication Software*: 04)
4. Teoría de la Señal y Comunicaciones (*Communication and Signal Theory*: 04):
 - Tratamiento de la Información (*Signal Processing*: 01)
 - Sistemas y Circuitos (*Circuits and Systems*: 02)
 - Campos Electromagnéticos (*Electromagnetic Fields*: 03)
 - Instrumentación (*Instrumentation*: 04)
5. Ciencia de los Materiales (*Materials Science*: 08):

- Materias para la Tecnología de la Información (*Materials for Information Technology*: 01)
- 6. Organización de Empresas (*Business Management*: 06):
 - Economía y Gestión de Empresas (*Business*: 01)
- 7. Ingeniería de Sistemas (*System Engineering*: 07):
 - Sistemas Concurrentes (*Concurrent Systems*: 01)
 - Ingeniería de Control (*Control Engineering*: 02)
- 8. Intensificación (*Suplement*: 08):
 - Sistemas y redes de telecomunicación (*Communication networks and systems*: 01)
 - Planificación y gestión de telecomunicaciones (*Communication planning and management*: 02)

Junto a estas áreas se encuentran Física Aplicada, Matemática Aplicada, Estadística e Investigación Operativa, las cuales no están incluidas en el corpus por manejar un lenguaje básico, herramienta necesaria para entender las demás asignaturas en las que se utiliza y de las que forma parte. Por ello se considera suficiente la proporción en la que aparece este lenguaje como parte de las materias restantes.

Las asignaturas denominadas *Ampliación*, *Complemento* o *Laboratorio* de una asignatura principal tampoco están incluidas, puesto que en la asignatura principal ya se encuentra el contenido básico con su lenguaje característico, para más tarde ampliar los conocimientos o poner en práctica la teoría. Éste es el caso, por ejemplo, de *Complementos de Telemática*, *Laboratorio de Electrónica*, *Laboratorio de Software de Comunicaciones*, etc. En la selección de las muestras se buscan aquellas que representen los contenidos de las dos partes de la materia y en el caso del laboratorio se introducen muestras de ejercicios prácticos.

Dentro del área de *Ingeniería Telemática*, se han agrupado bajo la etiqueta de Telemática (*Telematics*) varias asignaturas que tratan del funcionamiento de una red en su forma física: *Fundamentos de la Telemática*, *Telemática*, *Redes y Servicios de Comunicaciones*, *Sistemas de Telecomunicaciones* y *Conmutación*. De igual modo, en el área de *Teoría de la Señal y Comunicaciones*, se han agrupado bajo la etiqueta de *Tratamiento de la Información (Signal Processing)* las asignaturas que estudian el procesamiento de la señal: *Introducción a las Telecomunicaciones*, *Teoría de la Comunicación*, *Sistemas lineales*, *Comunicaciones Digitales*, *Tratamiento digital de la Señal y Transmisión de Datos*. El criterio de agrupación es fruto de la colaboración de personal especializado, profesores de las áreas de *Electrónica*, *Arquitectura y Tecnología de Computadoras*, *Telemática* y *Teoría de la Señal*. Aunque lo más

conveniente sería que el corpus fuera diseñado y construido por expertos, miembros de la comunidad discursiva, al menos en este trabajo se ha contado con la participación de muchos de ellos.

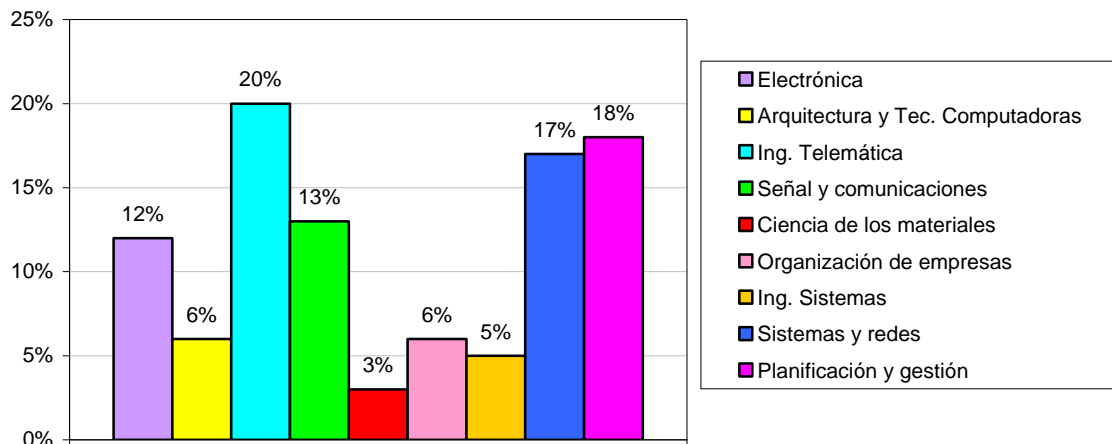


Figura 5.2 Volumen del corpus distribuido por áreas temáticas.

El último apartado contiene las asignaturas correspondientes a las dos intensificaciones disponibles en el ciclo largo de Telecomunicaciones, que permiten establecer un perfil de especialización dentro del título. La primera sección comprende la intensificación *Sistemas y Redes de Telecomunicación*, que compete a una especialización en el análisis de la señal; y la segunda, *Planificación y Gestión de Telecomunicaciones*, conducente a una especialización en telemática.

Finalmente, cada asignatura se compone de ocho divisiones que contienen los medios de difusión de donde proceden las muestras. Si se pretenden cubrir los tipos de texto o géneros característicos de la comunicación profesional y académica del sector, se deberían considerar las situaciones comunicativas como referentes para recoger muestras. Se estimarán solamente los actos de comunicación en los que al menos uno de los interlocutores sea un experto o profesional:

a) Comunicación entre expertos.

En este acto de comunicación, emisor y receptor son expertos en el tema. Se asume que los interlocutores, profesionales y/o docentes, comparten un alto nivel de conocimiento especializado de la materia en igual o similar grado. De la misma forma comparten un conocimiento del lenguaje especializado, así que utilizarán la terminología correspondiente, que le ha asignado un significado específico al lenguaje.

Este significado específico es definido previamente por un organismo externo para conseguir un estándar, es decir, para que exista un acuerdo común sobre el

significado único y explícito de un término cuando es utilizado por los expertos en el área.

En esta situación comunicativa probablemente aparezca una alta densidad de términos. Los textos característicos son los artículos de investigación, hojas de características, libros académicos, informes técnicos y documentos legales.

b) Comunicación de experto a iniciado.

El emisor posee un nivel más alto de conocimientos en la materia, de modo que se expresará de la misma forma que entre expertos pero, debido al tipo de audiencia, es probable que aclare o explique algunos términos que considere desconocidos o no comprendidos completamente por el receptor.

Esta situación se produce normalmente cuando el receptor pretende aumentar su nivel de conocimientos, tipos de receptor como son los estudiantes de últimos cursos, colegas con menos experiencia profesional o dedicados a aspectos diferentes de la misma área, etc. El tipo de texto característico es el libro de texto.

c) Comunicación de experto a principiante.

Los receptores se consideran principiantes o legos, si no poseen un conocimiento previo de la materia. Se trata de estudiantes que deben adquirir tales conocimientos con fines educativos o profesionales, y también receptores que no están relacionados directamente con la materia pero tienen cierto interés en ella.

El emisor utilizará una terminología apropiada o un lenguaje técnico simplificado, introduciendo aclaraciones con más frecuencia y explicándose en la lengua general, asumiendo siempre el bajo nivel de conocimientos del receptor. No obstante, el emisor requiere que el receptor aprenda y entienda el significado específico que se ha asignado al término utilizado en este dominio.

Los textos que se producen en esta situación son los artículos de revistas de carácter divulgativo, los manuales de instrucciones, los libros de textos de carácter introductorio, etc.

Las fuentes de donde proceden las muestras son las siguientes:

1. Revistas especializadas (*Magazines*: 01).

Artículos procedentes de revistas especializadas pero de carácter divulgativo.

2. Libros (*Books*: 02).

Textos procedentes de libros científicos y libros de texto, informes técnicos y las denominadas *tutorials*: lecciones o explicaciones difundidas por profesores, científicos, profesionales y empresas sobre un tema o área en un estilo formal.

3. Divulgación por internet (*Web*: 03).

Información obtenida de páginas web y cualquier tipo de texto disponible en internet y no clasificable dentro de los restantes apartados.

4. Artículos científicos (*Research*: 04).
Estudios dirigidos a especialistas de la materia con el fin de comunicar los resultados o el estado de una investigación: Artículos de investigación, artículos presentados en congresos y artículos publicados en revistas especializadas de carácter científico.
5. Resúmenes (*Abstracts*: 05).
Textos breves escritos que representan y preceden el contenido de un artículo, ponencia, etc.
6. Folletos (*Brochures*: 06).
Folletos que contienen información introductora o publicitaria; apartado donde se incluyen los manuales de instrucciones.
7. Anuncios, propaganda (*Advertising*: 07).
Anuncios publicitarios relacionados con cualquier aspecto de las telecomunicaciones.
8. Noticias tecnológicas (*Technology News*: 08).
Noticias relacionadas con cualquier aspecto de las telecomunicaciones.

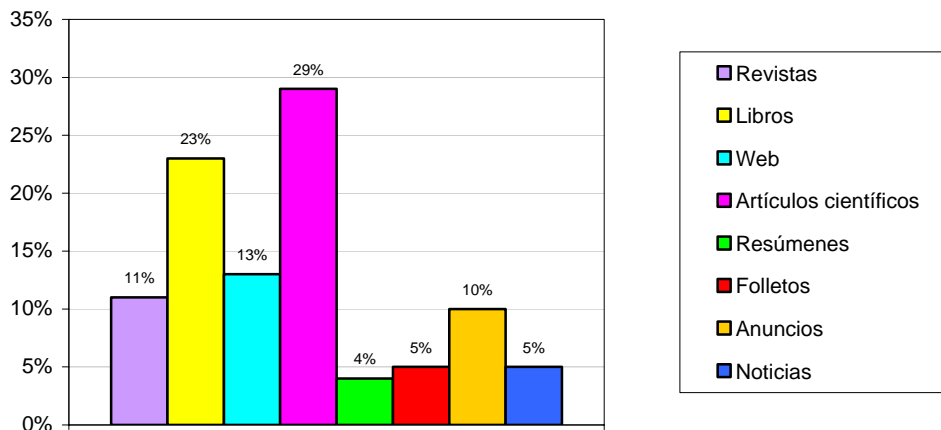


Figura 5.3 Volumen del corpus distribuido por tipos de texto.

El número o abreviatura que aparece al lado de cada sección forma parte de la etiqueta que compone el código de cada texto. De este modo, las muestras quedan archivadas con una estructura embebida y asociadas a un código que permitirá su posterior indexación. A modo de ejemplo, se presenta a continuación la estructura jerarquizada y el correspondiente código, de las muestras contenidas en el área de conocimiento de Arquitectura y Tecnología de Computadoras de procedencia británica:

<i>Telecommunications Corpus</i>
<i>British English BE</i>
<i>Computing Architecture and Technology 02</i>
<i>Computing Fundamentals 01</i>
<i>Magazines 01</i>
<i>Books 02</i>
<i>Web 03</i>
<i>Research 04</i>
<i>Abstracts 05</i>
<i>Brochures 06</i>
<i>Advertising 07</i>
<i>Technology News 08</i>
<i>Digital Electronic Systems 02</i>
<i>Magazines 01</i>
<i>Books 02</i>
<i>Web 03</i>
<i>Research 04</i>
<i>Abstracts 05</i>
<i>Brochures 06</i>
<i>Advertising 07</i>
<i>Technology News 08</i>

Cuadro 5.2 Código completo de ATC de procedencia británica

CAPÍTULO 6

APLICACIONES INFORMÁTICAS PARA EL PROCESAMIENTO DEL CORPUS

6.1 Software para el procesamiento del corpus.

El tipo de software más conocido aplicable a un corpus y disponible para su uso en ordenadores personales es el denominado concordanciero (*concordancing software* o *concordancers*).

Un concordanciero es un programa informático que se utiliza para identificar las ocurrencias de una palabra o cadena de palabras en un texto. Este tipo de software lee el texto línea a línea y proporciona de forma automática listados de frecuencia y líneas de concordancia de una palabra. Normalmente no está limitado a una cantidad máxima de texto o tamaño de archivo, su única restricción es la cantidad de espacio libre disponible en el disco duro y el tamaño y tiempo de acceso de la memoria RAM, de la que dependerá un procesamiento más rápido o más lento de la cantidad total de texto.

Las concordancias son un listado de todas las veces que aparece una palabra seleccionada en su contexto. El nodo, como se denomina genéricamente a la palabra seleccionada, aparece en la pantalla situado en el centro de la línea, uno debajo de otro y acotado hacia derecha e izquierda por un número determinado de las palabras que forman su co-texto.

Para que el programa sea capaz de procesar los textos, éstos deben encontrarse en formato ASCII (*American Standard Code for Information Interchange*), donde sólo aparecen las palabras, los signos de puntuación y los espacios en blanco. Así que se elimina cualquier otro tipo de información extraña al procesador de texto o restos de otros sistemas semióticos integrados en el texto original.

El concordanciero con el que se ha tratado el corpus de telecomunicaciones forma parte del paquete de aplicaciones de WordSmith Tools, adecuado para el volumen del corpus que se está manejando y que permite realizar un análisis tanto cuantitativo como cualitativo.

Este programa fue diseñado por Michael Scott especialmente para el manejo de corpus y ha sido ampliamente utilizado en estudios basados en corpus lingüísticos. El

propio diseñador define WordSmith como “*an integrated suite of programs for looking at how words behave in texts. You will be able to use tools to find out how words are used in your own texts*” (Scott, 1998). El conjunto de herramientas permiten procesar los textos de formas diferentes y facilita al usuario la identificación de patrones en el lenguaje.

Existe otro tipo de concordancieros denominados *Indexing concordancers* que reconocen el texto previamente indexado. El programa filtra las muestras según el índice propuesto, desplegando los datos procesados casi de forma instantánea.

Cada muestra en pantalla va encabezada por un código que cifra la información introducida previamente en la indexación del corpus. De este modo es posible conocer la procedencia y ubicación de las muestras, de acuerdo con los criterios de selección establecidos en la compilación. Por ejemplo, un artículo de una revista científica americana sobre electrónica digital iría precedido del siguiente código:

<A TCUS010301>

El indexador con el que se ha procesado el corpus de telecomunicaciones es un programa elaborado por el grupo de investigación LACELL de la Universidad de Murcia. El corpus indexado se analizará por medio de Tecnolingua, el programa asociado al indexador. En el Apéndice I¹ se encuentra el código completo utilizado en la indexación del corpus.

6.2 Información estadística básica sobre la composición del corpus.

Una vez que se ha conseguido el corpus apropiado, se procede al subsiguiente análisis. WordSmith aporta dos tipos de resultados inmediatos tras procesar el texto: Información estadística básica sobre la composición del corpus (Cuadro 6.1), y listados de las palabras que componen el corpus, ordenadas alfabéticamente y por orden descendente de la frecuencia en que ocurren.

El programa identifica 5.533.705 *tokens* definidos como secuencias de caracteres separadas por espacios en blanco o signos de puntuación. Naturalmente, no todas las palabras son distintas, muchas se repiten y con bastante frecuencia. El número de *types* o formas indica la cantidad de palabras diferentes que contiene el corpus, incluyendo diversas flexiones de verbos, nombres y adjetivos. El conjunto de formas constituye el vocabulario del texto. En la siguiente secuencia de cuatro *tokens* “*controls, controlled, controller, controls*”, aparecen tres formas distintas de un único lema: *control*. El concepto de lema se corresponde con el de voz o entrada léxica en un diccionario, es

¹ Todos los apéndices se encuentran en el CD adjunto.

decir, la palabra que representa de forma canónica todas las flexiones de la misma palabra.

La relación existente entre el número total de palabras y el número total formas viene dado por los cocientes *type/token ratio* y *standardised type/token*. Estos cocientes ofrecen una visión de la diversidad léxica del texto.

Type/token ratio se obtiene al dividir el total de palabras diferentes entre el número total de ocurrencias, multiplicado por 100. Cuanto más alto sea el resultado, mayor será la diversidad léxica de la muestra. Por el contrario, cuanto menor sea el resultado, mayor será la repetición y será necesario conocer menos formas para entender el texto. A medida que un texto crece, la probabilidad de encontrar palabras nuevas disminuye gradualmente.

El programa computa el *standardised type/token ratio* cada n palabras, siendo por defecto n=1.000. El cociente se calcula con los primeros 1.000 *tokens*, después se computa de nuevo para los siguientes 1.000, y así sucesivamente hasta el final del texto, presentando la media de los valores obtenidos. En el corpus de Telecomunicaciones, aparecen como media 38,26 palabras diferentes por cada secuencia de texto de 1.000 *tokens*.

La información estadística aporta igualmente datos sobre el número de oraciones y párrafos; longitud media de palabra, frase y párrafo; número de palabras según su número de letras, etc. Sin embargo, no muestra la cifra correspondiente al fenómeno denominado *hapax legomena*, es decir, las palabras que ocurren sólo una vez en todo el corpus. Aunque éstas se pueden identificar en el listado de frecuencias, parece oportuno incluir aquí este dato con el fin de completar la visión de conjunto sobre la composición del corpus.

Corpus de Telecomunicaciones	
<i>Tokens</i>	5.533.705
<i>Types</i>	59.826
<i>Type/token Ratio</i>	1,08
<i>Standardised Type/token</i>	38,26
<i>Sentences</i>	223.278
<i>Sentences length</i>	23,87
<i>Paragraphs</i>	30.472
<i>Paragraphs length</i>	102,95
<i>Hapax legomena</i>	21.755

Cuadro 6.1 Información estadística del Corpus de Telecomunicaciones.

Para identificar el comportamiento del lenguaje en un registro especializado es conveniente tener acceso a un corpus del lenguaje general, que actúe como base de comparación: “*Register analyses require a comparative approach: we need a baseline for comparison to know whether the use of a linguistic feature in a register is rare or common*” (Biber et al., 1998:136). De este modo, es posible contrastar el comportamiento lingüístico de los dos registros e identificar las características compartidas y distintivas de cada uno. La comparación de los resultados permite obtener una idea del grado de restricción de tales características en el dominio específico.

Igualmente, WordSmith requiere dos corpus de diferente tamaño para poder ejecutar algunas de sus herramientas. El programa opera utilizando como referencia el corpus mayor, con el cual compara el corpus de menor tamaño. En este estudio se ha utilizado el corpus de inglés general Lacell² (21 millones de palabras) como punto de referencia lingüístico. El corpus de referencia debe ser lo suficientemente grande para que aparezcan palabras de alta y baja incidencia. El Cuadro 6.2 muestra la información estadística básica sobre la composición del corpus Lacell.

Corpus Lacell	
<i>Tokens</i>	21.016.312
<i>Types</i>	176.237
<i>Type/token Ratio</i>	0,84
<i>Standardised Type/token</i>	42,76
<i>Sentences</i>	1.099.049
<i>Sentences length</i>	18,57
<i>Paragraphs</i>	240.146
<i>Paragraphs length</i>	181,60
<i>Hapax legomena</i>	77.969

Cuadro 6.2 Información estadística del Corpus Lacell.

La comparación de los valores obtenidos revela una mayor diversidad léxica en el corpus técnico que en el general (1,08 frente a 0,84 respectivamente). Este resultado se debe a la diferencia de tamaño y a la especialización del corpus de telecomunicaciones. Una muestra mayor incorpora más palabras repetidas y reduce la adquisición de palabras nuevas, puesto que el aumento de formas nuevas y de las frecuencias de las formas sigue la curva de una hipérbola (Figura 6.1) (Sánchez y Cantos, 1997). Por otro

² LACELL, corpus de 21 millones de palabras elaborado por el Grupo de Investigación LACELL, de la Universidad de Murcia. El corpus LACELL tiene la misma estructura que el corpus CUMBRE, de español (propiedad de SGEL, s.a.)

lado, es probable que un corpus técnico contenga más formas debido a la especialización temática donde los términos se utilizan con mayor afinidad.

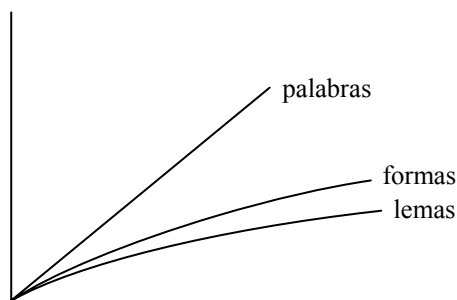


Figura 6.1 Relación entre el número de palabras, formas y lemas.

6.2.1 Listados de frecuencias: comparación general de los dos corpus.

Los listados de frecuencia presentan el tipo de información esencial que se puede extraer del lenguaje contenido en un corpus. Para obtener estos datos, el programa simplemente cuenta el número de veces que aparece una palabra y muestra el resultado en la pantalla del ordenador en cuestión de segundos.

En la Tabla 6.1 se encuentra el listado de las palabras más frecuentes en el corpus técnico y en el general. Debido a la gran cantidad de datos, se ha decidido mostrar sólo las 200 primeras a modo de ilustración. No obstante, los listados completos se hallan en el Apéndice I.

Teleco			Lacell			Teleco			Lacell		
1	THE	374.598	THE	1.264.637	101	NO	5.118	WAY	21.811		
2	OF	170.493	OF	621.663	102	MUST	5.087	MANY	21.695		
3	AND	145.104	AND	585.829	103	WOULD	5.072	WORK	21.187		
4	TO	140.361	TO	567.100	104	CURRENT	5.064	MUCH	20.807		
5	A	131.832	A	481.107	105	VALUE	5.063	RIGHT	20.779		
6	IN	104.696	IN	400.840	106	POWER	5.031	BACK	20.642		
7	IS	94.630	THAT	246.485	107	HOW	5.002	THOSE	20.602		
8	FOR	63.834	IS	234.895	108	TECHNOLOGY	4.969	SUCH	20.439		
9	THAT	52.034	IT	192.879	109	SHOULD	4.941	GO	20.420		
10	ARE	41.713	FOR	190.835	110	MANY	4.935	MAKE	20.343		
11	BE	41.543	I	166.771	111	YOUR	4.916	YEAR	20.144		
12	AS	40.644	ON	155.413	112	MANAGEMENT	4.890	YEARS	20.070		
13	THIS	37.780	YOU	153.066	113	ABOUT	4.879	GOOD	19.955		
14	WITH	36.388	WAS	147.651	114	BECAUSE	4.841	THREE	19.400		
15	ON	33.072	WITH	141.515	115	HOWEVER	4.827	US	19.188		
16	BY	31.017	BE	139.048	116	PROTOCOL	4.742	DOWN	18.902		
17	IT	28.218	AS	133.936	117	WHAT	4.616	BEING	18.871		
18	AN	27.199	ARE	115.124	118	TYPE	4.612	MADE	18.699		
19	CAN	25.288	AT	109.246	119	DO	4.596	EVEN	18.528		
20	OR	24.488	THIS	104.352	120	WORK	4.594	GOING	18.013		
21	FROM	21.405	HAVE	104.051	121	SOFTWARE	4.575	THROUGH	17.891		
22	AT	20.529	HE	101.706	122	FREQUENCY	4.551	DID	17.818		
23	WE	17.733	BY	100.557	123	INTERNET	4.504	GOT	17.668		
24	WHICH	17.453	BUT	97.716	124	USERS	4.493	THAT'S	17.568		

25	NOT	17.069	OR	90.550	125	APPLICATION	4.478	BEFORE	17.363
26	NETWORK	16.649	FROM	90.259	126	STATE	4.458	SAY	17.279
27	WILL	16.128	NOT	90.238	127	LAYER	4.425	LAST	16.970
28	HAVE	16.114	THEY	87.747	128	CASE	4.375	BETWEEN	16.489
29	DATA	14.613	WE	85.858	129	SINGLE	4.368	TAKE	16.455
30	HAS	13.444	AN	72.286	130	INPUT	4.347	DAY	16.306
31	ONE	13.218	HIS	71.362	131	TRAFFIC	4.345	HERE	15.877
32	SYSTEM	12.624	WHICH	71.192	132	MOBILE	4.341	USED	15.508
33	TIME	12.391	ONE	69.030	133	POINT	4.236	OFF	15.083
34	USED	11.874	ALL	64.772	134	OUTPUT	4.139	YES	15.028
35	IF	11.826	WILL	64.656	135	OUT	4.129	STILL	14.873
36	ALL	11.602	THERE	63.786	136	ORDER	4.120	USE	14.866
37	THESE	11.577	HAD	62.946	137	PROVIDE	4.119	OWN	14.833
38	MORE	11.454	IF	60.795	138	WELL	4.096	TOO	14.490
39	YOU	11.448	SO	58.064	139	WIRELESS	4.083	LONG	14.487
40	ALSO	10.382	HAS	57.865	140	SUPPORT	4.077	I'M	14.370
41	OTHER	10.291	THEIR	55.555	141	NEED	4.013	COME	14.201
42	USE	10.255	WERE	54.699	142	WHILE	4.010	LITTLE	14.183
43	EACH	10.224	WHAT	53.429	143	ADDRESS	3.951	EACH	14.172
44	SUCH	10.006	CAN	53.245	144	CIRCUIT	3.932	MUST	13.847
45	ITS	9.714	ABOUT	50.694	145	SECTION	3.912	OLD	13.647
46	WHEN	9.681	MORE	48.002	146	ROUTER	3.910	LIFE	13.644
47	WAS	9.664	BEEN	47.960	147	END	3.872	BOTH	13.561
48	SYSTEMS	9.479	WOULD	47.699	148	LINK	3.853	WANT	13.515
49	TWO	9.407	WHEN	47.304	149	VERY	3.835	SAME	13.512
50	USING	9.214	UP	46.684	150	C	3.819	WORLD	13.292
51	INFORMATION	9.161	WHO	44.477	151	RESULTS	3.712	ANOTHER	13.291
52	BUT	8.933	DO	44.356	152	FIG	3.702	NEED	13.073
53	THEIR	8.930	YOUR	44.275	153	AREA	3.680	PART	12.903
54	THEY	8.734	NO	44.264	154	WITHIN	3.677	HOME	12.880
55	BASED	8.448	OUT	44.010	155	RESEARCH	3.672	MR	12.659
56	BETWEEN	8.403	SHE	41.286	156	AVAILABLE	3.627	AGAIN	12.545
57	NEW	8.347	HER	40.851	157	B	3.609	PUT	12.491
58	THAN	8.052	SOME	39.255	158	WAY	3.599	THINGS	12.205
59	I	8.004	SAID	39.028	159	DIGITAL	3.595	UNDER	12.188
60	ONLY	7.939	TIME	38.072	160	SECURITY	3.578	FOUR	12.182
61	MAY	7.808	THEM	36.684	161	PACKET	3.577	HOUSE	12.179
62	NUMBER	7.759	TWO	35.842	162	SERVER	3.574	MIGHT	12.132
63	DESIGN	7.701	OTHER	34.595	163	SEE	3.552	REALLY	12.024
64	THERE	7.494	ER	34.335	164	THREE	3.545	LOOK	11.988
65	FIGURE	7.325	LIKE	34.311	165	ROUTING	3.542	SOMETHING	11.888
66	INTO	7.308	IT'S	33.956	166	THEM	3.531	COURSE	11.872
67	BEEN	7.150	WELL	33.887	167	INTERFACE	3.526	GREAT	11.862
68	CONTROL	7.124	PEOPLE	33.784	168	SHOWN	3.507	WHILE	11.765
69	SERVICE	7.085	MY	33.455	169	BIT	3.479	MAN	11.736
70	SIGNAL	7.022	INTO	33.322	170	RATE	3.467	DOES	11.586
71	SOME	6.904	THEN	32.603	171	COULD	3.466	NEVER	11.438
72	E	6.607	VERY	31.790	172	SINCE	3.452	NUMBER	11.407
73	EXAMPLE	6.379	ITS	31.536	173	DEVICES	3.430	SAYS	11.193
74	HIGH	6.348	NOW	31.441	174	LIKE	3.415	NEXT	11.121
75	S	6.315	THAN	30.916	175	N	3.408	DIFFERENT	10.974
76	FIRST	6.314	JUST	30.398	176	LOW	3.404	PLACE	10.885
77	USER	6.292	ONLY	30.211	177	FUNCTION	3.380	GIVE	10.867
78	DIFFERENT	6.198	NEW	29.580	178	GIVEN	3.368	FIVE	10.727

79	SERVICES	6.161	THINK	29.358	179	FOLLOWING	3.361	MEAN	10.692
80	ANY	6.052	ANY	28.742	180	DOES	3.338	ALWAYS	10.688
81	ACCESS	5.999	KNOW	28.695	181	PROBLEM	3.273	END	10.678
82	PROCESS	5.949	THESE	27.710	182	PROVIDES	3.257	WHY	10.617
83	OVER	5.925	MAY	27.626	183	POSSIBLE	3.221	EVERY	10.580
84	SAME	5.904	FIRST	27.523	184	COST	3.214	VOICE	10.536
85	THEN	5.897	COULD	27.493	185	CHANNEL	3.212	AGAINST	10.495
86	MODEL	5.895	OUR	26.860	186	FIELD	3.205	CHILDREN	10.400
87	SO	5.861	ALSO	26.777	187	MAKE	3.200	GOVERNMENT	10.202
88	NETWORKS	5.832	ME	26.279	188	NOW	3.164	YEAH	10.150
89	SET	5.813	OVER	26.142	189	BEING	3.162	WOMEN	10.149
90	PERFORMANCE	5.686	HOW	25.149	190	COMMUNICATION	3.159	HOWEVER	10.078
91	UP	5.652	BECAUSE	25.110	191	COMPUTER	3.156	FIND	10.064
92	MOST	5.602	HIM	25.062	192	ANALYSIS	3.153	POINT	10.057
93	WHERE	5.463	GET	24.367	193	COMMUNICATIONS	3.144	HIGH	9.997
94	APPLICATIONS	5.414	SEE	23.623	194	SOURCE	3.120	AROUND	9.932
95	WERE	5.378	ERM	23.162	195	BANDWIDTH	3.119	INFORMATION	9.807
96	BOTH	5.355	AFTER	23.041	196	CODE	3.112	HELP	9.788
97	LEVEL	5.309	MOST	22.716	197	STANDARD	3.081	WITHOUT	9.641
98	IP	5.239	WHERE	22.333	198	MULTIPLE	3.046	PER	9.569
99	THROUGH	5.202	SHOULD	22.216	199	APPROACH	3.043	SINCE	9.536
100	OUR	5.152	DON'T	22.022	200	DEVELOPMENT	3.032	FEW	9.526

Tabla 6.1 Listado de las 200 palabras más frecuentes en el corpus técnico y en el general.

Los listados de frecuencia han aportado conclusiones muy útiles sobre el funcionamiento del vocabulario. Una de las más importantes es el hecho de que las palabras más frecuentes cubren un gran porcentaje de ocurrencias en una lengua (Sinclair, 1991; Schmitt, 2000). Como se puede apreciar en la Tabla 6.1, *the* es la palabra más frecuente en los dos corpus, ocupando un 6,77% del total de palabras en el técnico y un 6,02% en el general. Sinclair, en el análisis estadístico del corpus Cobuild, concluye que más de la mitad de un texto es producto de la repetición de palabras (Sinclair, 1991).

En inglés general, las 3 palabras más frecuentes forman un 11% del total, las 10 más frecuentes un 22%, las 50 más frecuentes un 37%, las 100 más frecuentes un 44% y las 2.000 más frecuentes alrededor del 80% (Schmitt, 2000:73). Estas cifras concuerdan con los resultados obtenidos en los corpus utilizados en este estudio, produciéndose tan sólo unas leves diferencias:

Palabras más frecuentes	% en Lacell	% en Teleco
3 cubren 11%	11	12
10 - 22%	21	23
50 - 37%	38	36
100 - 44%	46	42
2.000 - 80%	78	79

Cuadro 6.3 Proporción de corpus que cubren las palabras más frecuentes.

De la muestra de 5,5 millones de palabras del corpus específico, solamente 59.826 son formas diferentes; de éstas, 21.755 (36%) aparecen una vez en el texto, constituyendo sólo el 0,39% del total. Más de 34.000 palabras aparecen de 1 a 3 veces, las que se repiten 100 veces apenas pasan de 30 y sin embargo, alrededor de 750 palabras se utilizan más de 1.000 veces en el total de las muestras. En definitiva, más de la mitad del texto se forma por medio de la repetición de palabras.

Los listados de frecuencia han revelado que las palabras de mayor incidencia suelen ser gramaticales o funcionales. Este grupo está compuesto por verbos auxiliares, verbos modales, pronombres, artículos, preposiciones y conjunciones. Tales palabras contribuyen esencialmente a formar la estructura gramatical del lenguaje, no confieren significado léxico y su comportamiento no varía.

Frente a esta categoría se encuentran las palabras nocionales o de contenido, que transmiten la mayor parte del contenido léxico. Sustantivos, adjetivos, verbos y adverbios forman este grupo. Al contrario que las palabras funcionales, las palabras nocionales dependen de la variedad del lenguaje registrado en el corpus.

La distinción entre palabras nocionales y funcionales permite medir la complejidad léxica de un texto por medio del índice de densidad léxica, es decir, la proporción de palabras nocionales expresada en un porcentaje. Para calcularlo, el número total de palabras de contenido se divide entre el total de *tokens* y se multiplica por cien. El índice de densidad léxica será más alto en los textos que contienen una proporción mayor de palabras nocionales. De acuerdo con los estudios realizados por Ure (Ure, 1971 en Stubbs, 2001), los textos escritos tienden a presentar una densidad léxica del 40%, oscilando entre el 36 y 57%. Tras realizar el cálculo correspondiente, se observa que el corpus de telecomunicaciones obtiene un valor que se aproxima al nivel superior de densidad léxica:

Corpus Teleco	
Palabras de contenido	3.076.453
Total <i>tokens</i>	5.533.705
Densidad léxica	55,59%

Cuadro 6.4 Densidad léxica

Las palabras más frecuentes tienden a guardar una distribución estable, de modo que cualquier cambio notable en el orden puede ser significativo (Sinclair, 1991:31). En un corpus general, alrededor de las 100 palabras más frecuentes son funcionales. Por lo tanto, la intrusión de palabras nocionales en este rango señala un comportamiento destacable.

En la Tabla 6.1, se puede observar que la primera palabra de contenido en Lacell es *said*, que alcanza la posición 59. En el corpus Teleco, la primera palabra de contenido se adelanta a la posición 26: *network*. En corpus especializados, las 50 palabras más frecuentes son sustancialmente diferentes a las 50 más frecuentes de un corpus general (Kennedy, 1998:101). Las 6 palabras más frecuentes en los dos corpus coinciden completamente: *the, of, and, to, a, is*. A continuación, siguen apareciendo palabras funcionales, pero en distinto orden. En el corpus técnico, se produce una alternancia de palabras funcionales y de contenido a partir de la posición 26 (*network*) hasta la 68 (*control*), donde empiezan a predominar las nocionales. En Lacell, a su vez, la alternancia se produce en la franja que se extiende desde la posición 101 (*way*) hasta la 157 (*put*). Cuanto más especializado es un corpus, más palabras de contenido aparecen en los niveles de alta frecuencia, mientras que en corpus generales empiezan a predominar a partir de las 150 palabras más frecuentes (Kennedy, 1998:102).

La selección de las 50 palabras nocionales más frecuentes en ambos corpus manifiesta un contenido léxico de naturaleza distinta (Tabla 6.5). En el corpus de Teleco, el contenido está relacionado con la disciplina representada, mientras que en Lacell, el léxico es de carácter general. Además, para reunir esta cantidad de palabras, hay que llegar a niveles de frecuencia distintos en los dos corpus. En el técnico, es posible extraer 25 palabras nocionales de las 81 más frecuentes y las restantes 25, de las 127 más frecuentes. En el corpus general es necesario recorrer una franja mayor para realizar la misma operación. Las 25 primeras palabras nocionales se pueden localizar entre las 139 palabras más frecuentes y las siguientes 25 alcanzan la posición 184.

Igualmente, existe una gran diferencia entre la cantidad de palabras de contenido que aparecen entre las 100 más frecuentes. En Lacell sólo 8 se encuentran en este nivel, frente a las 36 identificadas en Teleco. Estas últimas ocurren con una frecuencia de entre 5.514 y 16.649, cubriendo respectivamente entre 0,10% y 0,30% del total del corpus.

La mayor presencia de palabras nocionales en los niveles de alta frecuencia del corpus técnico puede ser indicativa de su densidad léxica. Además, si este hecho se considera junto al carácter especializado de las 50 palabras nocionales más frecuentes (Tabla 6.2), se podría esperar una mayor presencia de términos específicos en el corpus de Teleco que en el general.

	Teleco	Lacell		Teleco	Lacell
1	NETWORK	SAID	26	PROCESS	LONG
2	DATA	TIME	27	MODEL	LIFE
3	SYSTEM	LIKE	28	NETWORKS	OLD
4	TIME	WELL	29	SET	COME
5	USED	PEOPLE	30	PERFORMANCE	WANT
6	USE	NEW	31	WHERE	WORLD
7	SYSTEMS	THINK	32	APPLICATIONS	NEED
8	USING	KNOW	33	BOTH	PART

9	INFORMATION	WAY	34	LEVEL	HOME
10	BASED	WORK	35	IP	PUT
11	NEW	RIGHT	36	CURRENT	THINGS
12	ONLY	GO	37	VALUE	HOUSE
13	NUMBER	MAKE	38	POWER	REALLY
14	DESIGN	YEAR	39	TECHNOLOGY	LOOK
15	FIGURE	YEARS	40	MANAGEMENT	COURSE
16	CONTROL	MADE	41	PROTOCOL	GREAT
17	SERVICE	GOING	42	TYPE	MAN
18	SIGNAL	DID	43	WORK	DOES
19	EXAMPLE	GOT	44	SOFTWARE	NUMBER
20	HIGH	SAY	45	FREQUENCY	SAYS
21	FIRST	LAST	46	INTERNET	DIFFERENT
22	USER	TAKE	47	USERS	PLACE
23	DIFFERENT	DAY	48	APPLICATION	GIVE
24	SERVICES	USE	49	STATE	MEAN
25	ACCESS	USED	50	LAYER	VOICE

Tabla 6.2 Las 50 palabras de contenido más frecuentes.

El corpus de Teleco por sí mismo, aunque contenga más muestras de palabras especializadas que el corpus general, no puede revelar si tales palabras caracterizan el lenguaje representado. El índice de frecuencia no constituye un parámetro que determine si una palabra es típica del registro. Para ello, es necesario comparar las frecuencias de las palabras en el registro técnico y en el general. No obstante, este tipo de información cuantitativa proporciona una base empírica idónea para realizar un estudio cualitativo. Se podría examinar, por ejemplo, la distribución de una palabra de alta frecuencia a lo largo del discurso, qué palabras suelen co-ocurrir, etc... El comportamiento estadístico de una palabra supone el punto de partida para su análisis contextual.

7.1 Justificación del estudio del vocabulario técnico desde una perspectiva didáctica.

El criterio más general que define el carácter especializado de un lenguaje, está basado en la temática y la terminología utilizada para denominar conceptos y transmitir información. Fundamentalmente, son los especialistas de cada materia quienes conocen los términos y los manejan de forma natural en el discurso.

Si el objetivo perseguido es la comunicación efectiva en un ámbito científico-técnico, será preciso averiguar qué léxico caracteriza esta variedad. Por lo tanto, el motivo para distinguir el vocabulario técnico de cualquier otro tipo de vocabulario, responde a la necesidad de identificar las palabras que serán especialmente útiles a estudiantes de un lenguaje específico. Sobre una base de conocimiento del inglés como lengua general, sería provechoso aprender el inglés específico que se utilizará en el entorno académico y futuro profesional.

Aunque en el lenguaje existe una gran cantidad de palabras diferentes, no todas son igual de útiles. Una forma de medir la utilidad de una palabra es considerar con qué frecuencia ocurre. Entre las palabras altamente incidentes se encuentran las funcionales y un grupo numeroso de nocionales. Si el estudiante conoce estas palabras, será capaz de reconocer una gran proporción en un texto y alcanzará un buen grado de comprensión.

En el lenguaje general, un vocabulario formado por las 2.000 palabras más frecuentes permite entender el 80% de las palabras en un texto (Nation y Waring, 1997). En un dominio específico, se debería obtener una base formada por las palabras más frecuentes, más el vocabulario específico del área. Nation (2001:187) afirma al respecto: “*When learners have mastered the 2,000-3,000 word of general usefulness in English, it is wise to direct vocabulary of learning to more specialised areas*”. Las palabras que ocurren fuera de este nivel de alta incidencia se utilizan principalmente en situaciones muy limitadas y específicas. De este modo, tendría sentido concentrarse en el vocabulario de frecuencia más baja que aparece con más probabilidad en el área de interés.

Según Yang (1986), la proporción de texto que cubren las 2.000 palabras más frecuentes, sólo explica hasta qué punto las palabras del texto resultan familiares al lector y no significa que se haya comprendido el mensaje. La alta concentración de términos en el lenguaje específico entorpece la comprensión del texto. Por lo tanto, el alumno debe haber adquirido un cierto número de términos científico-técnicos para conseguir una competencia lectora práctica. Con este argumento, Yang defiende la inclusión de tales términos en la enseñanza del inglés técnico.

El estudio del vocabulario técnico tiene sentido en un contexto de aprendizaje de la asignatura o área correspondiente. De acuerdo con Nation (2001), el objetivo principal para identificar los diferentes tipos de vocabulario es aportar una base sólida para planificar la enseñanza y aprendizaje del lenguaje, centrándose en las palabras que aparecen con frecuencia y que presentan una alta distribución.

Nation sostiene igualmente que, cuando una palabra presenta dos acepciones, una de alta frecuencia y otra de uso técnico, se deberían estudiar las conexiones y diferencias entre los dos significados, siendo el profesor el encargado de preparar a los alumnos para enfrentarse y manejar tales términos.

7.2 Clasificación de los diferentes tipos de vocabulario en un texto especializado por medio de criterios cualitativos.

Gran parte de la literatura relacionada con el inglés para fines específicos considera la presencia de diferentes tipos de vocabulario en un texto especializado y recoge la clasificación correspondiente. Aunque las descripciones realizadas por medio de análisis cualitativos han intentado definir este fenómeno de forma detallada, no se ha conseguido establecer una definición exacta de cada categoría, ni un método efectivo para su clasificación. Los límites que separan un grupo de otro no están totalmente definidos y a veces se solapan. Además, la terminología adoptada varía ligeramente entre autores.

Kennedy y Bolitho (1984) mencionan cinco categorías para poder organizar la enseñanza del vocabulario específico, aunque no especifican claramente el método a seguir para identificar y catalogar los diferentes tipos de vocabulario. Las categorías son las siguientes:

1. Abreviaciones técnicas.
2. Símbolos y fórmulas.
3. Vocabulario altamente técnico (*Highly technical*).
4. Vocabulario subtécnico, formado por las palabras que no son específicas de la materia de la especialidad (*subject specific*), pero que aparecen regularmente en

textos científicos y técnicos. Estas palabras son normalmente conocidas en el inglés general, pero adquieren un significado especial en un contexto científico o técnico.

5. Otro tipo de vocabulario especializado. Estaría compuesto por el vocabulario específico característico de una materia, por ejemplo: palabras con raíces latinas y griegas del inglés jurídico; neologismos, acrónimos y acuñaciones en informática, etc.

A diferencia de las cinco categorías de Kennedy y Bolitho, Robinson (1991) hace referencia a numerosos autores (Martin, 1976; Inman, 1978; Loots, 1986; Baker, 1988; King, 1989; Hoffman, 1986; etc...) que identifican tres niveles de vocabulario en un texto especializado y que son conscientes de su importancia en la enseñanza de lenguas.

El primer nivel consiste en el vocabulario especializado o terminología, pero no ofrece ninguna definición ni forma de identificarlo. El segundo nivel adquiere varias denominaciones, tales como: vocabulario semitécnico, vocabulario subtécnico o vocabulario general científico-técnico. Algunos autores identifican este nivel como vocabulario académico cuando lo caracterizan por comprender las palabras comunes a diferentes áreas científicas o técnicas. Finalmente, el tercer nivel está compuesto por el vocabulario general, no académico.

Dudley-Evans y St. John (1998) parten de la aceptación general de la importancia de la enseñanza del vocabulario en el inglés con fines específicos. Consideran el vocabulario semitécnico, subtécnico o *core vocabulary*, como el vocabulario común en un campo específico. Reconocen que no existe una definición satisfactoria de este concepto y aluden a la clasificación de Baker (1988), quien diferencia seis categorías distintas de vocabulario (Dudley-Evans, 1998:82):

1. *Items which express notions general to all specialised disciplines;*
2. *general language items that have a specialised meaning in one or more disciplines;*
3. *specialised items that have different meanings in different disciplines;*
4. *general language items that have a restricted meaning in different disciplines;*
5. *general language items that are used to describe or comment on technical processes or functions in preference to other items with the same meaning, for example occur rather than happen;*
6. *items used to signal the writer's intentions or evaluation of material presented.*

Dudley-Evans y St. John identifican principalmente dos áreas que agrupan las seis categorías propuestas por Baker. La primera comprende las categorías 1, 5 y 6,

compuesta por el vocabulario utilizado en la lengua general, que aparece con mayor frecuencia en un campo específico. El término adoptado es vocabulario semitécnico o *core vocabulary*. Por otro lado, las categorías 2, 3 y 4 abarcan el vocabulario provisto de un significado restringido especializado y que puede variar de unas disciplinas a otras. Los usos especializados del vocabulario general en una disciplina específica forman el vocabulario técnico. Por último, mencionan que el vocabulario semitécnico se puede identificar por medio de corpus lingüísticos informatizados pero no especifican cómo, e igualmente citan a varios autores que aportan listados de palabras.

Finalmente, Alcaraz Varó (2000) clarifica la diferencia entre vocabulario y léxico. Aunque los dos comprenden en principio la totalidad de las palabras o unidades léxicas de una lengua, el léxico es una clase abierta que se enriquece constantemente, y el vocabulario es una clase cerrada (vocabulario de un autor, de una especialidad). El vocabulario de las lenguas de especialidad se denomina de las siguientes formas: metalenguaje, terminología o tecnolecto (toda lengua de especialidad), y suele ordenarse en tres grupos (Alcaraz, 2000:42-44):

1. Terminología o vocabulario técnico.

La terminología está formada por las unidades léxicas de carácter técnico, llamadas ‘términos’, cuyos significados están definidos de forma unívoca dentro de una teoría (Cabré, 1993). Los términos son monosémicos, no necesitan del contexto para activar su significado, pues se entienden dentro de una teoría o de un campo del saber. La terminología presenta menor dificultad a los estudiantes de lenguas de especialidad por varias razones, entre las que destacan: a) su monosemia y precisión de su definición, b) su carácter medular, ya que, si no se entiende, no se puede comprender la especialidad, y c) su fácil homologación de una lengua a otra.

Algunos ejemplos son: *azufre, software, meningitis, etc.*

2. El vocabulario semitécnico o subtécnico.

Está formado por unidades léxicas del lenguaje común que han adquirido uno o varios significados nuevos dentro de un campo del saber; es lo que Sager et al. (1980) llaman “unidades léxicas generales re-denominadas” (*re-designated general language items*).

Este vocabulario es polisémico, se ha formado en la mayoría de los casos por extensión del significado mediante el proceso de analogía, añadiendo acepciones adicionales al significado adicional. Es muy abundante en cualquier especialidad, suele ser bastante problemático porque sus acepciones en el léxico general inadvertidamente se transfieren al léxico de la especialidad.

Algunos ejemplos son: *defence, strength, beam, force, etc.*

3. El vocabulario general de uso frecuente en una especialidad.

Está formado por las palabras del léxico general, que sin perder su significado propio, viven dentro o en los aledaños de la especialidad. No son palabras técnicas en el sentido estricto del término, porque conservan su significado primitivo, pero por su elevado índice de presencia en una especialidad son imprescindibles.

Algunos ejemplos son: *test, repair, performance, prove, check*, etc.

7.3 Clasificación de los diferentes tipos de vocabulario en un texto especializado por medio de criterios cuantitativos y corpus lingüísticos.

La relación entre la lingüística del corpus y el inglés con fines específicos ha dado lugar a una serie de estudios relacionados con la identificación y clasificación del vocabulario presente en un texto especializado. Los criterios cualitativos se complementan y combinan con criterios cuantitativos con la intención de facilitar la identificación de términos. La ventaja de aplicar análisis cuantitativos radica en que se puede obtener información estadísticamente significativa. A los fenómenos lingüísticos observados se les asigna un índice de frecuencia que los cuantifica y que permite realizar cálculos estadísticos y probabilísticos.

Yang Huizhong (1986) propone una técnica para identificar términos científico-técnicos y caracterizar tipos de textos de forma automática por medio de un algoritmo. Se centra en dos tipos básicos de términos: términos de una palabra (*single-word terms: hardware, terminal*, etc.) identificados por su frecuencia y distribución; y términos multipalabra (*multi-word terms: programming language, radio wave transmitter*, etc.), identificados por sus colocaciones, es decir, cuando dos palabras o más co-ocurren frecuentemente.

Según su hipótesis, los términos científico-técnicos deberían presentar una frecuencia alta al ser susceptibles al tema, a la vez que su frecuencia varía en materias distintas. Por lo tanto, sería posible identificar los términos científico-técnicos por su comportamiento estadístico. Los siguientes parámetros establecen las diferentes categorías: Desviación estándar relativa, frecuencia media, distribución, *peakratio* y *rangeratio*.

En estos términos distingue palabras subtécnicas, es decir, palabras de contenido, no funcionales, que tienen una alta distribución, pero una frecuencia relativamente baja. Presentan nociones generales que comparten la mayoría de las áreas científicas. Su desviación estándar relativa, *peakratio* y *rangeratio* son bastante bajas. Por otro lado, los términos científico-técnicos presentan una distribución muy baja, pero su *peakratio*

y *rangeratio* son muy altos, los cuales señalan la especificidad del campo científico-técnico.

Persiguiendo un objetivo distinto al anterior, se halla Farrell (1990), que pretende identificar el lenguaje general de la ciencia (*specific common core*). Éste se tomaría como referencia para desarrollar un programa léxico, común a varias ramas de inglés científico-técnico. Asimismo, busca una definición más precisa de vocabulario semitécnico, mediante la aplicación de criterios cuantitativos y cualitativos. Utilizando un corpus de 20.000 palabras, estudia el vocabulario del inglés de la electrónica por medio de la frecuencia y distribución de palabras, y compara el listado de palabras semitécnicas de la electrónica con la lista de vocabulario semitécnico específico general (*common core*) que obtiene aplicando sus criterios al corpus LOB.

Farrell revisa los criterios que aportan varios autores para identificar el vocabulario semitécnico y poder alcanzar una conclusión común. Todas las definiciones comparten un criterio básico: el vocabulario semitécnico aparece en diferentes áreas científicas. Por otro lado, se incluye que las palabras de esta categoría se definen independientemente del contexto y no son técnicas. El factor frecuencia va unido a diferentes posibilidades de distribución:

1. Si la frecuencia y distribución de una palabra son altas, esta palabra formaría parte del vocabulario semitécnico.
2. Si la frecuencia de una palabra es alta, pero su distribución es baja, esta palabra formaría parte del vocabulario técnico. Las palabras de alta frecuencia que claramente no forman parte del *common core* de las ciencias, pero que son características de la disciplina en particular, suelen presentar una distribución baja.
3. A veces una palabra presenta frecuencia y distribución altas, pero pertenece al vocabulario técnico, por ejemplo: *electrical*. En este caso, se aplica un criterio cualitativo para su identificación.

El vocabulario semitécnico no aparece normalmente en cursos de lengua general, de modo que habría que eliminar las palabras básicas o generales más las gramaticales para poder identificarlo. Normalmente se toman como referencia las palabras que aparecen en *General Service List* de West (1953). Sin embargo, Farrell utiliza un método subjetivo para definir el vocabulario general: palabras que probablemente aparecen en un curso de inglés general de nivel intermedio bajo.

Farrell aplica todos los criterios de selección al corpus LOB, elimina las palabras generales, las específicas de la materia, las de baja frecuencia y baja distribución, y halla así las palabras que formarían el vocabulario semitécnico o vocabulario general de

la ciencia. En definitiva, como resumen de su estudio, obtiene las siguientes conclusiones:

1. Las palabras generales son aquellas ya conocidas por los alumnos de un curso de inglés específico.
2. Las palabras técnicas son términos puramente técnicos, desconocidos en el lenguaje general, por ejemplo: *anode, diode, filament*, etc., o palabras que ocurren generalmente pero que tienen un significado técnico preciso, por ejemplo: *energy, force, wave, signal*, etc.
3. Las palabras técnicas de alta frecuencia se pueden identificar objetivamente por su baja distribución. El hecho de que una palabra técnica presente baja frecuencia y baja distribución muestra que está restringida a un campo más estrecho dentro del área de la electrónica, por ejemplo: *transistor, collector, silicon*.

Finalmente, el vocabulario semitécnico se define por comprender palabras formales, independientes del contexto, que ocurren con una frecuencia y distribución alta en diferentes disciplinas científicas, y que no aparecen normalmente en los cursos de inglés general.

En la presente revisión cronológica, se introduce un nuevo tipo de vocabulario, que forma parte de los usos del lenguaje técnico específico: el registro académico. Coxhead (2000) evalúa y reforma la lista de vocabulario académico que elaboró en 1998. El corpus utilizado contiene alrededor de 3,5 millones de palabras procedentes de textos académicos de distintas disciplinas. Para que una familia de palabras (un lema, sus flexiones y otras palabras relacionadas morfológicamente) sea seleccionada como parte del vocabulario académico, debe cumplir los siguientes requisitos:

1. *Specialised occurrence*: La familia de palabras tiene que estar fuera de las primeras 2.000 palabras que ocurren con más frecuencia en la *General Service List* de West.
2. Distribución: Un miembro de la familia de palabras tiene que ocurrir al menos 10 veces en cada una de las cuatro secciones del corpus y en 15 o más de las 28 áreas temáticas.
3. Frecuencia: Los miembros de una familia de palabras tienen que ocurrir al menos 100 veces en el corpus académico.

El resultado del estudio da lugar a un total de 570 familias de palabras que formarían el vocabulario académico. Este listado demuestra ser especialmente útil en el lenguaje de la ciencia ya que consigue cubrir el 9,1% del subcorpus analizado. Además, entre otros resultados, el estudio indica que la mayoría de las palabras en la lista de

vocabulario académico está asociada particularmente con la expresión escrita académica.

El listado de vocabulario académico está incluido en la clasificación de Nation (2001), distinguiendo cuatro tipos de vocabulario en un texto: palabras de alta frecuencia, palabras académicas, palabras técnicas y palabras de baja frecuencia.

1. Palabras de alta frecuencia. Este tipo de vocabulario está compuesto por palabras funcionales y palabras de contenido que aparecen en la lista de West (1953) *A General Service List of English Words*, compuesta por alrededor de 2.000 familias de palabras. Las primeras 1.000 cubren el 77% de las palabras de un texto y las 1.000 restantes alrededor del 5%. Estas palabras presentan una alta frecuencia y alta distribución.
2. Palabras académicas. Palabras que aparecen en la lista de Coxhead (1998) *Academic Word List*, que contiene 570 familias de palabras y que normalmente suponen un 9% de las palabras de un texto académico.
En algunas ocasiones se ha denominado vocabulario académico al vocabulario subtécnico porque no contiene palabras técnicas sino más bien palabras formales.
3. Palabras técnicas. Palabras de contenido estrechamente relacionadas con el tema específico del texto, de modo que varían según el área; o bien palabras de alta frecuencia que tienen un significado especializado. Estas palabras son razonablemente comunes en un campo del saber, pero no tan comunes fuera de él. Normalmente cubren alrededor del 5% de las palabras del texto.
4. Palabras de baja frecuencia. Cubren un 5% de las palabras de un texto académico. Existen miles de ellas en la lengua, es el grupo de palabras más grande. Incluye todas las palabras que no son de alta frecuencia, ni académicas, ni técnicas de un área específica. Este grupo está formado por palabras técnicas de otros campos, nombres propios, palabras que casi entran en la lista de palabras de alta frecuencia y palabras que raramente aparecen en el uso diario del lenguaje.

Después de hacer esta distinción, Nation divide el vocabulario de los textos académicos en tres niveles: *general service* o vocabulario básico, vocabulario subtécnico o académico y vocabulario técnico. Los dos primeros niveles los identifica con las listas de palabras de West y Coxhead respectivamente. Dentro del vocabulario técnico establece cuatro categorías, dependiendo de la frecuencia relativa de la forma y el significado de una palabra técnica. Ésta se caracteriza por pertenecer específicamente a un tema, campo o disciplina.

- Categoría 1: La palabra raramente aparece fuera del campo específico, si es que aparece, siendo exclusiva del área tanto en forma como en significado. Sólo el estudio de la materia determinada permite una comprensión total, por lo tanto se trata de palabras claramente técnicas.
Ejemplos en informática: *wysiwyg, rom, pixel, modem*; en electrónica: *Anode, impedance, galvanometer, dielectric*.
- Categoría 2: La palabra se utiliza dentro y fuera del área en particular, pero no con el mismo significado. Al igual que la categoría anterior, son claramente palabras técnicas, pues su significado más general no lleva de manera directa a su significado en el contexto técnico.
Ejemplos en informática: *execute, scroll, paste*; en electrónica: *induced, flux, terminal, earth*.
- Categoría 3: La palabra se utiliza tanto dentro como fuera del área en particular, pero la mayoría de sus usos con significado específico, aunque no todos, ocurren en este campo. Además, este significado específico es accesible directamente por medio de su significado en un contexto general fuera del área.
Ejemplos en informática: *memory, drag, window*; en electrónica: *coil, energy, positive, gate, resistance*.
- Categoría 4: La palabra es más común en este campo que en cualquier otro contexto. Existe poco o ningún significado específico, aunque alguien con conocimientos sobre el tema tendría una idea más precisa de lo que quiere decir.
Ejemplos en informática: *print, program, icon*; en electrónica: *drain, filament, load, plate*.

Los niveles de tecnicidad varían según del grado de restricción de una palabra a un área. Las palabras que pertenecen a las categorías 3 y 4 no son únicas de un determinado campo ni en forma ni en significado, de modo que su clasificación como técnicas es menos obvia. Para distinguir las palabras que pertenecen a las categorías 2 y 3 es necesario considerar no sólo la forma sino también el significado.

De acuerdo con Nation, una forma sistemática de elaborar una lista de vocabulario técnico sería por medio de listados de frecuencia basados en corpus lingüísticos. Se podría utilizar un listado que aporte cifras de la frecuencia y distribución de palabras en un corpus general, y compararlo con el correspondiente listado de un corpus específico. Sin embargo, Nation no proporciona los datos necesarios para realizar tal análisis.

En los siguientes puntos, Nation sintetiza las aportaciones de varios autores que utilizan un corpus especializado para clasificar el vocabulario:

1. Un grupo pequeño de palabras altamente relacionadas con el campo específico ocurren con una frecuencia muy alta, cubriendo al menos un 10% de todas las palabras del corpus. Por ejemplo: *current, fig., voltage, circuit*.
2. Varias palabras de alta frecuencia aparecen también en la lista de West con significados relacionados. No son palabras completamente técnicas, pero también son conocidas en contextos no especializados. Por ejemplo: *energy, plate, supply, positive, flow*.
3. Algunas palabras de alta incidencia destacan por reflejar la naturaleza del discurso, por ejemplo: *fig., we, if*.
4. Varias palabras claramente técnicas en forma y significado presentan una frecuencia muy baja en el corpus, lo que dificulta su distinción sólo por medio del criterio de frecuencia.
5. Algunas palabras técnicas tienen una distribución muy baja dentro del corpus específico. Esto significa que las ocurrencias de una palabra técnica determinada no se extienden homogéneamente en un texto, sino que tienden a agruparse en un capítulo o sección.

En la misma línea que el trabajo anterior, Chung y Nation (2003) describen su estudio realizado sobre el vocabulario técnico de anatomía y lingüística aplicada, comparando los listados de frecuencia del corpus específico con los de un corpus general. El objetivo principal es comprobar si el vocabulario técnico se puede identificar de forma fiable y obtener alguna indicación sobre el tamaño y la densidad del vocabulario técnico.

De acuerdo con los autores, las investigaciones realizadas sobre el vocabulario técnico han subestimado la función de este vocabulario en los textos especializados y además, existe una falta de información sobre cómo se relaciona el vocabulario técnico con otros tipos de vocabulario.

En este estudio se sigue la clasificación del vocabulario propuesta por Nation (2001): palabras de alta frecuencia, palabras de baja frecuencia, vocabulario académico o semitécnico y vocabulario técnico. Este último comprende tanto palabras técnicas que pueden aparecer en usos no especializados del lenguaje general, como palabras técnicas únicas de un campo especializado en particular.

Tras procesar el corpus específico, al listado de frecuencias obtenido le sustraen las palabras que aparecen en *Academic Word List* de Coxhead y las palabras de la lista *General Service* de West. El resto de palabras comprendería el vocabulario técnico. Utilizan también criterios cualitativos para identificar las palabras generales que puedan tener un uso específico en el campo que se está tratando.

En un segundo trabajo, Chung (2003) identifica los términos técnicos del lenguaje de anatomía, comparando la frecuencia de las palabras en un corpus específico con un corpus general. La autora parte de los siguientes supuestos:

1. La frecuencia de las palabras no técnicas (*non-term*) debería ser aproximadamente la misma en el corpus específico y en el general.
2. La frecuencia de los términos técnicos debería ser mucho más alta en el corpus específico que en el general.

Tras realizar las operaciones pertinentes, decide tomar el ratio 50 como valor mínimo que distingue a una palabra como término (vocabulario técnico). Es decir, cuando una palabra es aproximadamente 50 veces más frecuente en el corpus específico que en el general se selecciona como término.

La razón por la que toma esta cifra como ratio de corte deriva de la comparación de los resultados con los obtenidos en la clasificación experimental realizada previamente. En la denominada *Rating Scale Approach*, un experto en la materia identifica cualitativamente los términos específicos. Tras comparar los resultados, las palabras con ratio > 50 coinciden con los términos identificados por el experto, con una fiabilidad del 86%.

7.4 Combinación de criterios.

Con la intención de aunar criterios y según las referencias y aportaciones anteriores, a continuación se reflejan todos los puntos considerados para identificar y definir cada tipo de vocabulario. Un texto especializado, en este caso textos de la ingeniería de telecomunicación, presenta cuatro tipos de vocabulario: general, académico, técnico y semitécnico.

7.4.1 Vocabulario general.

La definición de este tipo de vocabulario es la más clara. Este grupo estaría compuesto por palabras funcionales o gramaticales, más las palabras de contenido más frecuentes en la lengua general. Normalmente, la lista de 2.000 palabras elaborada por West (1953), *General Service List of English Words* es aceptada como referencia.

Las palabras que pertenecen a esta categoría suelen aparecer en los cursos de inglés general de nivel intermedio bajo.

Respecto a su comportamiento estadístico, estas palabras se pueden distinguir por su alta frecuencia y alta distribución. Suelen corresponderse con las 2.000 palabras que ocurren con más frecuencia en un corpus general de un millón de palabras.

7.4.2 Vocabulario académico.

El vocabulario académico está formado por las 570 familias de palabras que aparecen en la lista elaborada por Coxhead y revisada en el 2000, *Academic Word List*.

Estas palabras se caracterizan por ser razonablemente frecuentes en textos académicos de un amplio repertorio de disciplinas académicas; son relativamente infrecuentes en otros tipos de textos como novelas o en el lenguaje oral coloquial; proceden principalmente del latín, francés y griego; y no están relacionadas con ningún tema en especial (Wang y Nation, 2004).

7.4.3 Vocabulario técnico.

El vocabulario técnico se compone de las palabras de contenido estrechamente relacionadas con la disciplina o tema específico, que caracterizan el lenguaje de la disciplina como un área individual de la lengua y que componen la terminología específica. Las palabras puramente técnicas son desconocidas en el lenguaje general y parte de ellas proceden del griego y del latín.

Los estudios consultados coinciden en que los términos específicos presentan una frecuencia alta en comparación con su frecuencia en un corpus general, a la vez que su distribución es estrecha en el corpus específico (Yang, 1986; Farell, 1990; Nation, 2001; Chung, 2003). Dependiendo del estudio, algunos autores afirman que los términos específicos pueden identificarse también según los siguientes parámetros:

- *Peakratio* y *rangeratio* muy altos (Yang, 1986).
- Su frecuencia en un texto específico es al menos 50 veces mayor que su frecuencia en un texto general (Chung, 2003).

Dentro del vocabulario técnico existen palabras que también se utilizan en la lengua general, pero que adquieren un significado específico diferente dentro de una disciplina. Su significado general no conduce a su significado correspondiente en el contexto técnico. Algunos autores consideran este grupo de palabras como vocabulario semitécnico (Sager et al., 1980; Alcaraz, 2001).

Finalmente se manifiesta un grupo de palabras que cualitativamente son reconocidas como términos técnicos, pero que cuantitativamente presentan, o frecuencia y distribución altas, o bien frecuencia y distribución bajas. Este último comportamiento indica que el término pertenece a un campo más estrecho dentro del área específica que se trate.

A continuación, se aproximan los límites difusos que separan o diferencian el vocabulario técnico del semitécnico y el semitécnico del académico. Una vez que se han identificado las palabras propias del área específica, se aprecia un grupo de palabras

que, por su significado y comportamiento estadístico, algunos autores catalogan como vocabulario semitécnico o académico y otros como vocabulario técnico de distinto grado de especificidad.

7.4.4 Vocabulario semitécnico.

Este tipo de vocabulario adquiere varias denominaciones: vocabulario subtécnico, vocabulario general de la ciencia (*specific common core*), vocabulario procedimental (*procedural vocabulary*) y vocabulario académico, aunque dependiendo del autor, a veces el concepto varía ligeramente.

El vocabulario semitécnico está compuesto por palabras de contenido, no específicas de la especialidad, que adquieren un significado especial en un contexto científico-técnico. Dicho significado especial es accesible directamente por medio de su significado en un contexto general. En ocasiones, este grupo se considera técnico por ser utilizado especialmente en un área específica, pero se le otorga un menor grado de especificidad.

Esta categoría también se define por contener palabras formales, no técnicas, independientes del contexto y que aparecen en distintos campos técnicos y académicos con más frecuencia que en un texto específico. Tal perspectiva lleva a confundir los límites entre el vocabulario semitécnico y el académico, aunque es posible que una palabra pertenezca a los dos grupos a la vez y las categorías se solapen.

Siguiendo criterios cuantitativos, este grupo se caracteriza por presentar una alta distribución y su frecuencia es mayor en un corpus específico que en uno general.

7.5 Métodos para seleccionar los tipos vocabulario.

Resulta evidente que el soporte informático es imprescindible para aplicar un método de selección de vocabulario, que exige: procesar los corpus, manejar los listados de palabras, identificar el comportamiento estadístico de las palabras y poder emplear tanto criterios cuantitativos como criterios cualitativos. Con este soporte, se pueden aplicar principalmente dos métodos de selección: método de listas de corte y método de comparación de corpus.

7.5.1 Método de listas de corte.

Este es el método más rudimentario, cuya aplicación requiere tan sólo un corpus específico y listas de corte. Una lista de corte contiene aquellas palabras que no se desea que aparezcan en el listado de frecuencia de un corpus. En los estudios citados anteriormente, las listas de corte utilizadas contienen palabras generales, funcionales y académicas, las cuales se han sustraído del corpus específico. Para tal fin se suelen

emplear la lista de West (1953), *General Service List of English Words* y la de Coxhead (2000), *Academic Word List*.

El resultado de la sustracción debería contener el vocabulario técnico con sus distintos grados de especificidad, o en otras palabras, el vocabulario técnico y semitécnico. El inconveniente de utilizar este método reside en que requiere la aplicación de más criterios para cribar los datos derivados y afinar la selección de palabras. Normalmente, este método supone un procedimiento base para reducir primero el volumen de datos que se manejan, y después aplicar una combinación de criterios cualitativos y cuantitativos.

7.5.2 Método de comparación de corpus.

Este método está basado en la comparación estadística de la frecuencia de las palabras en un corpus específico y otro general. El primer paso consiste en obtener los listados de frecuencia de cada corpus y después, existen dos opciones. Por un lado, se eliminan las palabras que no aparecen en los dos corpus, guardando sólo las comunes y, a continuación, se normalizan los resultados para poder proceder a su comparación. La segunda opción requiere normalizar la frecuencia de todas las palabras desde un principio, otorgándole valor 0 a las palabras que no aparecen en alguno de los corpus. Con estos resultados se pueden seguir los diferentes criterios de clasificación, conforme a la información obtenida de los estudios consultados y reflejados a lo largo de este capítulo.

Una vez que el comportamiento de las palabras adquiere valores numéricos, es posible aplicar tests estadísticos u otros algoritmos más sofisticados (Oakes, 1998; Peñas et al., 2001; Moore, 2002) que permiten obtener una mayor afinidad y fiabilidad en los resultados. Estas fórmulas suelen tener en cuenta no sólo la frecuencia sino también la distribución de las palabras en varios niveles. La distribución adquiere diferentes valores según se considere la presencia de una palabra en los diferentes textos, secciones, subcorpus o corpus operativos.

Los métodos de comparación basados en la frecuencia y distribución son los más utilizados para la identificación de términos dado que, generalmente, éstos son más frecuentes en los textos especializados que en los generales, la característica estadística típica de los términos. En cualquier caso, la constante supervisión de los resultados impera en la aplicación de todo tipo de criterio, ya que la clasificación automática de palabras es de tipo morfológico. Los programas informáticos sólo reconocen secuencias de caracteres, no distinguen significados y por lo tanto, no son capaces de realizar una clasificación semántica de las palabras.

ANÁLISIS DEL CORPUS Y RESULTADOS

El análisis del corpus técnico pretende dar respuesta a las preguntas de investigación formuladas en relación con los objetivos específicos planteados y, de este modo, alcanzar la consecución del objetivo principal de la investigación.

La caracterización del perfil léxico de las telecomunicaciones está enfocada desde dos perspectivas diferentes y en función de dos parámetros: el grado de restricción y el grado de relevancia o representatividad de una palabra en el ámbito específico.

El primer enfoque está orientado hacia la clasificación del vocabulario, de acuerdo con los diferentes niveles del léxico presentes en un lenguaje de especialidad. Con el fin de llevar a cabo dicha taxonomía, se seguirán dos procedimientos. En primer lugar, se utilizará el método de las listas de corte para identificar el vocabulario específico del corpus y los términos únicos de cada área de conocimiento. Mediante la aplicación de un segundo método, que contrasta el corpus específico y el corpus general, se tratará de agrupar los diferentes tipos de vocabulario adoptando los criterios propuestos por Chung (2003). De esta manera se pretende contestar la siguiente pregunta: ¿Qué tipos de palabras se encuentran en el inglés de las telecomunicaciones?

La segunda perspectiva cambia el plano de visión y persigue una caracterización léxica atendiendo al grado de relevancia de las palabras dentro del registro, independientemente del nivel al que pertenezcan. El procedimiento que se va a seguir consiste en la aplicación de un test estadístico que discrimina las palabras clave, es decir, aquellas palabras que son estadísticamente significativas. Esta operación permitirá responder a la pregunta: ¿Qué palabras son clave o representativas del inglés de la ingeniería de telecomunicación?

8.1 Clasificación del vocabulario: Aplicación del método de listas de corte.

8.1.1 Listas de corte y filtrado del corpus.

La primera aproximación consiste en la aplicación de listas de corte, como se ha explicado anteriormente en el apartado 7.5.1, con el fin de identificar el vocabulario

específico de las telecomunicaciones. Este proceso consta de dos partes: La elección de las listas de corte y el filtrado del corpus completo.

8.1.1.1 Elección de las listas de corte.

Las listas de corte utilizadas para filtrar el corpus específico contienen palabras generales, funcionales y académicas. Estos grupos incluyen las 2.000 palabras más frecuentes de la lengua inglesa, procedentes de la *General Service List of English Words* (West, 1953), más el vocabulario académico de la *Academic Word List*. (Coxhead, 2000).

Las primeras mil palabras más frecuentes comprenden aquellas cuya frecuencia es superior a 332 por cada 5 millones de palabras, más los meses del año, días de la semana, números, tratamientos (*Mr, Mrs, Miss, Ms, Mister*) y saludos frecuentes (*Hello, Hi, etc*). Las listas de corte están formadas por familias de palabras (Bauer y Nation, 1993), cubriendo el lema base, sus diferentes formas y los restantes miembros de la familia. Por ejemplo *assume*, va acompañada de *assumed, assumes, assuming, assumption* y *assumptions*. Además incluyen las palabras con doble ortografía en inglés americano y británico, como *organise* y *organize*.

Estas listas se encuentran disponibles y preparadas en un formato adecuado dispuestas para su utilización en la página web www3.vuw.ac.nz/lals/staff/paul-nation/nation.aspx, por gentileza de Paul Nation.

No obstante, es imprescindible revisar los listados para identificar las palabras generales o académicas que adquieren un significado especializado dentro del campo de las telecomunicaciones. Si tales palabras permanecieran en el listado de corte, se estaría perdiendo información relativa a la especialidad. No sería apropiado eliminar del corpus específico palabras como *network, signal* o *system*, ya que forman parte del nombre de las mismas áreas temáticas que integran el corpus (*signal processing*) o incluso de una rama de especialización en los estudios de telecomunicación (*communication networks and systems*). Por lo tanto, se han extraído de las listas de corte, las siguientes familias de palabras estrechamente relacionadas con las telecomunicaciones: *access, assemble, bond, channel, code, communicate, component, compound, compute, concurrent, couple, convert, data, design, device, discrete, distribute, image, input, link, layer, logic, media, network, offset, output, overlap, process, protocol, route, simulate, signal, system, technology, transmit, transfer* y *transform*.

Tras revisar y unificar las listas de corte, quedan un total de 10.773 formas diferentes que se excluirán del corpus específico.

8.1.1.2 Filtrado del corpus completo.

El programa WordSmith permite introducir las listas de corte deseadas y filtrar el corpus. Como resultado se obtiene el listado de palabras ordenadas de forma descendente por la frecuencia en la que ocurren (Apéndice II). Las 59.826 formas que componen el total del corpus se reducen a 50.864 tras el filtrado. Sin embargo, la lista resultante precisa una segunda criba, que se realiza de forma manual, con el fin de retirar los siguientes tipos de palabras:

- Palabras en idiomas distintos al inglés: *asuntos, attaché, universidad, vivísimo, voilà*, etc.
- Nombres propios y apellidos: *Acevedo, Acharya, Martínez, Stuber, Sugymoto, Sumitaka, Zare, Zinio, Zutphen*, etc.
- Gentilicios y topónimos: *Cartagena, Chinese, French, Galway, Italy, Portugal, Spanish*, etc.
- Erratas o palabras con faltas de ortografía: *acheived, actuall, ajustement, aftern, amplidude, therefore, trafic, utput, variate*, etc.
- Palabras que se han juntado por error: *actiontooutputs, actorsmay, adevoted, alwaysunloaddii, aproblem, theswire, todigital, topeer*, etc.

Para las palabras con un apóstrofe señal del genitivo sajón o de la forma contraída de un verbo se han seguido dos caminos. Si la palabra pertenece a uno de los casos anteriores o a las listas de corte, se elimina directamente, de lo contrario, permanece en la lista. Por ejemplo se suprimen *Switzerland's, today's, let's, what's*, etc., y se mantienen *network's, router's, system's, software's*, etc.

Las palabras generales y académicas incluidas en las listas de corte cubren alrededor del 15% del total de las formas del corpus técnico. Los datos muestran que el corpus contiene una gran proporción de palabras (el 85%) que no aparece en el vocabulario más frecuente de la lengua general ni en el vocabulario académico.

El resultado de la sustracción debería contener el vocabulario técnico con sus distintos grados de especificidad o, en otras palabras, el vocabulario técnico y semitécnico. Sin embargo, como se señaló anteriormente, este método presenta el gran inconveniente de incluir entre los resultados un elevado número de palabras que no son verdaderos términos y palabras generales utilizadas en una gran variedad de temas. Por lo tanto, la aplicación de listas de corte, no se considera un método válido que permita clasificar adecuadamente el léxico en sus correspondientes niveles. En consecuencia, se decide tomar la lista resultante como listado de referencia del corpus de telecomunicaciones (Listado de referencia de Teleco), puesto que se consigue disminuir considerablemente el volumen de palabras original. Las 59.826 formas iniciales quedan

reducidas a 36.077. Sobre este listado, se aplicarán otros criterios de selección y se seguirán varios procedimientos para analizar el comportamiento léxico.

8.1.2 Distribución de las palabras por áreas de conocimiento.

Aunque la aplicación de las listas de corte no ha aportado los resultados deseados, contribuye positivamente a los propósitos de la investigación. El empleo de listas de corte resulta de utilidad para estudiar la distribución de las palabras por áreas de conocimiento. Este tipo de información permite identificar tanto las palabras que comparten todas las áreas, como las únicas propias de cada una de ellas. El procedimiento es ligeramente distinto al anterior e implica el filtrado del corpus por áreas de conocimiento, y el contraste de los resultados con el listado de referencia y con las áreas entre sí.

8.1.2.1 Filtrado del corpus por áreas de conocimiento.

En esta ocasión, el corpus se fragmenta en nueve archivos, de acuerdo con las áreas de conocimiento y las especialidades que lo integran. A continuación, se filtra cada archivo individualmente con la misma lista de corte utilizada para el corpus completo. De esta forma, se obtienen nueve listados independientes, preparados para el contraste.

8.1.2.2 Contraste de los listados.

En una hoja de Excel se introduce primero el listado de referencia de Teleco con el cual se va a cotejar cada una de las listas filtradas restantes. Tras las operaciones requeridas, se consigue plasmar en la pantalla del ordenador de forma gráfica qué palabra de la lista de referencia aparece o no aparece en cada una de las áreas temáticas. Si una palabra de la lista de referencia se halla también en un área, se marca con el número correspondiente al área de conocimiento, de lo contrario, se introduce un guión. La siguiente tabla ilustra el procedimiento realizado y muestra la distribución de las 150 palabras más frecuentes. En el Apéndice II se halla la tabla completa correspondiente al resultado del contraste, tanto el listado de referencia filtrado por completo (3) como el listado original de 50.864 palabras (4).

Telecom.	FREC.	Electr	Arq.Comp	Telmat	Señal	Materl	Empres	Sist	Esp.Señ	Esp.Telm
1 NETWORK	16649	1	2	3	4	5	6	7	801	802
2 DATA	14613	1	2	3	4	5	6	7	801	802
3 SYSTEM	12624	1	2	3	4	5	6	7	801	802
4 SYSTEMS	9479	1	2	3	4	5	6	7	801	802
5 DESIGN	7701	1	2	3	4	5	6	7	801	802
6 ACCESS	5999	1	2	3	4	5	6	7	801	802
7 PROCESS	5949	1	2	3	4	5	6	7	801	802
8 NETWORKS	5832	1	2	3	4	5	6	7	801	802
9 IP	5239	1	2	3	4	-	6	7	801	802
10 TECHNOLOGY	4969	1	2	3	4	5	6	7	801	802
11 PROTOCOL	4742	1	2	3	4	5	6	7	801	802

12	SOFTWARE	4575	1	2	3	4	5	6	7	801	802
13	INTERNET	4531	1	2	3	4	5	6	7	801	802
14	LAYER	4425	1	2	3	4	5	6	7	801	802
15	INPUT	4347	1	2	3	4	5	6	7	801	802
16	TRAFFIC	4345	1	2	3	4	5	6	7	801	802
17	MOBILE	4341	1	2	3	4	5	6	7	801	802
18	OUTPUT	4139	1	2	3	4	5	6	7	801	802
19	WIRELESS	4083	1	2	3	4	5	6	-	801	802
20	CIRCUIT	3932	1	2	3	4	5	6	7	801	802
21	ROUTER	3910	1	2	3	4	-	6	7	801	802
22	LINK	3905	1	2	3	4	5	6	7	801	802
23	FIG	3702	1	2	3	4	5	6	7	801	802
24	DIGITAL	3595	1	2	3	4	5	6	7	801	802
25	PACKET	3625	1	2	3	4	-	6	7	801	802
26	SERVER	3649	1	2	3	4	5	6	7	801	802
27	INTERFACE	3526	1	2	3	4	5	6	7	801	802
28	DEVICES	3430	1	2	3	4	5	6	7	801	802
29	CHANNEL	3212	1	2	3	4	5	6	7	801	802
30	COMMUNICATION	3159	1	2	3	4	5	6	7	801	802
31	COMPUTER	3201	1	2	3	4	5	6	7	801	802
32	COMMUNICATIONS	3144	1	2	3	4	5	6	7	801	802
33	BANDWIDTH	3119	1	2	3	4	5	6	7	801	802
34	CODE	3112	1	2	3	4	5	6	7	801	802
35	WEB	2978	1	2	3	4	5	6	7	801	802
36	VOLTAGE	2945	1	2	3	4	5	6	7	801	802
37	NODE	2853	1	2	3	4	5	6	7	801	802
38	OPTICAL	2822	1	2	3	4	5	6	7	801	802
39	SIMULATION	2817	1	2	3	4	5	6	7	801	802
40	DEVICE	2849	1	2	3	4	5	6	7	801	802
41	ALGORITHM	2799	1	2	3	4	5	6	7	801	802
42	PROCESSING	2770	1	2	3	4	5	6	7	801	802
43	COMPONENTS	2727	1	2	3	4	5	6	7	801	802
44	FILTER	2659	1	2	3	4	5	6	7	801	802
45	ARCHITECTURE	2581	1	2	3	4	5	6	7	801	802
46	TRANSMISSION	2544	1	2	3	4	5	6	7	801	802
47	MEMORY	2458	1	2	3	4	5	6	7	801	802
48	NODES	2361	1	2	3	4	5	6	7	801	802
49	PACKETS	2308	1	2	3	4	5	6	7	801	802
50	HARDWARE	2257	1	2	3	4	5	6	7	801	802
51	TECHNOLOGIES	2137	1	2	3	4	5	6	7	801	802
52	CIRCUITS	2122	1	2	3	4	5	6	7	801	802
53	CELL	2089	1	2	3	4	5	6	7	801	802
54	SWITCH	2075	1	2	3	4	5	6	7	801	802
55	PROTOCOLS	1996	1	2	3	4	-	6	7	801	802
56	COMPONENT	2000	1	2	3	4	5	6	7	801	802
57	LOGIC	1920	1	2	3	4	5	6	7	801	802
58	PROCESSES	1912	1	2	3	4	5	6	7	801	802
59	ETHERNET	1911	1	2	3	4	-	6	7	801	802
60	ROUTERS	1891	1	2	3	4	-	6	7	801	802
61	CONFIGURATION	1885	1	2	3	4	5	6	7	801	802
62	FIBER	1869	1	-	3	4	5	6	7	801	802
63	DISTRIBUTED	1824	1	2	3	4	5	6	7	801	802
64	DESIGNED	1822	1	2	3	4	5	6	7	801	802
65	ALGORITHMS	1777	1	2	3	4	5	6	7	801	802

66	REMOTE	1770	1	2	3	4	5	6	7	801	802
67	LOOP	1766	1	2	3	4	5	6	7	801	802
68	TCP	1717	1	2	3	4	-	6	7	801	802
69	CABLE	1715	1	2	3	4	5	6	7	801	802
70	ELECTRONIC	1698	1	2	3	4	5	6	7	801	802
71	ATM	1639	1	2	3	4	5	6	7	801	802
72	CLIENT	1601	1	2	3	4	5	6	7	801	802
73	STORAGE	1591	1	2	3	4	5	6	7	801	802
74	LINEAR	1590	1	2	3	4	5	6	7	801	802
75	CHANNELS	1569	1	2	3	4	5	6	7	801	802
76	VIDEO	1559	1	2	3	4	5	6	7	801	802
77	REFERENCE	1543	1	2	3	4	5	6	7	801	802
78	SPECTRUM	1499	1	2	3	4	5	6	7	801	802
79	LAN	1481	1	2	3	4	-	6	7	801	802
80	DATABASE	1451	1	2	3	4	5	6	7	801	802
81	DISTRIBUTION	1448	1	2	3	4	5	6	7	801	802
82	SATELLITE	1401	1	2	3	4	5	6	7	801	802
83	LINKS	1394	1	2	3	4	5	6	7	801	802
84	PORT	1392	1	2	3	4	5	6	7	801	802
85	WAVELENGTH	1352	1	2	3	4	5	6	7	801	802
86	JAVA	1365	1	2	3	4	5	6	7	801	802
87	MEDIA	1322	1	2	3	4	5	6	7	801	802
88	SWITCHING	1315	1	2	3	4	5	6	7	801	802
89	DESTINATION	1311	1	2	3	4	-	6	7	801	802
90	AL	1299	1	2	3	4	5	6	7	801	802
91	OSPF	1323	1	-	3	4	-	-	-	801	802
92	ANALOG	1264	1	2	3	4	5	6	7	801	802
93	MULTI	1260	1	2	3	4	5	6	7	801	802
94	TRANSFER	1254	1	2	3	4	5	6	7	801	802
95	DIRECTORY	1251	1	2	3	4	-	6	7	801	802
96	ANTENNA	1242	1	2	3	4	5	6	7	801	802
97	CHIP	1229	1	2	3	4	5	6	7	801	802
98	COMPUTING	1227	1	2	3	4	5	6	7	801	802
99	SERVERS	1208	1	2	3	4	5	6	7	801	802
100	FILTERS	1207	1	2	3	4	5	6	7	801	802
101	TRANSPORT	1195	1	2	3	4	5	6	7	801	802
102	INTERFACES	1187	1	2	3	4	5	6	7	801	802
103	EXECUTION	1186	1	2	3	4	5	6	7	801	802
104	QOS	1155	1	-	3	4	-	6	7	801	802
105	VHDL	1150	1	2	3	4	-	-	-	801	802
106	LINUX	1128	1	2	3	4	-	6	7	801	802
107	ET	1116	1	2	3	4	5	6	7	801	802
108	MPLS	1112	1	-	3	4	-	6	-	801	802
109	PC	1142	1	2	3	4	5	6	7	801	802
110	GSM	1109	1	-	3	4	-	6	-	801	802
111	SWITCHES	1107	1	2	3	4	-	6	7	801	802
112	TELECOMMUNICATIONS	1105	1	2	3	4	5	6	7	801	802
113	CARRIER	1145	1	2	3	4	5	6	7	801	802
114	PROCESSOR	1093	1	2	3	4	5	6	7	801	802
115	AUTHENTICATION	1082	1	-	3	4	-	6	7	801	802
116	ETC	1081	1	2	3	4	5	6	7	801	802
117	COMPUTERS	1076	1	2	3	4	5	6	7	801	802
118	SYNTHESIS	1054	1	2	3	4	5	6	7	801	802
119	BROADBAND	1049	1	2	3	4	5	6	7	801	802

120	IMAGE	1047	1	2	3	4	5	6	7	801	802
121	NETWORKING	1030	1	2	3	4	-	6	7	801	802
122	GRAPH	1016	1	2	3	4	5	-	7	801	802
123	VPN	1007	1	2	3	4	-	6	-	801	802
124	VENDORS	1004	1	2	3	4	5	6	7	801	802
125	IEEE	1002	1	2	3	4	5	6	7	801	802
126	CELLULAR	987	1	2	3	4	5	6	7	801	802
127	CELLS	957	1	2	3	4	5	6	7	801	802
128	TRANSMITTED	957	1	2	3	4	5	6	7	801	802
129	VECTOR	955	1	2	3	4	5	-	7	801	802
130	ENTERPRISE	953	1	2	3	4	5	6	7	801	802
131	DISCRETE	949	1	2	3	4	5	6	7	801	802
132	ONLINE	937	1	2	3	4	5	6	7	801	802
133	PLATFORM	935	1	2	3	4	5	6	7	801	802
134	RELAY	926	1	2	3	4	-	6	7	801	802
135	MODULATION	908	1	2	3	4	5	-	7	801	802
136	DSP	906	1	2	3	4	-	6	7	801	802
137	HEADER	903	1	2	3	4	5	6	7	801	802
138	SILICON	897	1	2	3	4	5	6	7	801	802
139	COM	890	1	2	3	4	5	6	7	801	802
140	FREQUENCIES	885	1	2	3	4	5	6	7	801	802
141	PROCESSORS	884	1	2	3	4	5	6	7	801	802
142	FEEDBACK	883	1	2	3	4	5	6	7	801	802
143	VENDOR	859	1	2	3	4	-	6	7	801	802
144	ARRAY	870	1	2	3	4	5	6	7	801	802
145	MAC	859	1	2	3	4	-	-	7	801	802
146	LSAS	858	-	-	3	-	-	-	-	-	802
147	DEFAULT	854	1	2	3	4	-	6	7	801	802
148	AMPLIFIER	851	1	2	-	4	5	6	7	801	802
149	FUZZY	851	1	2	3	4	-	6	7	801	802
150	MICROWAVE	848	1	2	3	4	5	6	7	801	802

Tabla 8.1 Distribución de las palabras por áreas de conocimiento.

A simple vista se puede observar que las palabras de frecuencia más alta suelen aparecer en todas o en la mayoría de las áreas. Por lo tanto, habrá que desplazarse hacia niveles de frecuencia inferiores para identificar aquellas palabras que sólo se manifiestan en un área de conocimiento. Asimismo, por la distribución de las palabras de alta frecuencia, se puede inferir que las áreas de conocimiento están estrechamente relacionadas, siendo el área de materiales (05) la que se distingue por compartir menos palabras. En total, 661 formas, de diferentes niveles de frecuencia (desde 14 hasta 16.649), aparecen en todas las áreas de conocimiento. De las palabras que permanecen tras los filtrados previos, un 1,8% es común a todas las áreas. La siguiente tabla muestra las 661 palabras con su correspondiente frecuencia.

Palabra	Frec.	Palabra	Frec.	Palabra	Frec.
NETWORK	16649	KEYWORDS	417	COMPRESSED	132
DATA	14613	SENSOR	416	EQUILIBRIUM	132
SYSTEM	12624	SIMULATED	413	PEAKS	131
SYSTEMS	9479	TECHNOLOGICAL	407	VERSA	131
DESIGN	7701	NONLINEAR	402	ACTIVATED	128
ACCESS	5999	OBJECTIVES	401	ROUTINES	128

PROCESS	5949	DIAL	397	OFFS	127
NETWORKS	5832	ANNOUNCED	394	CANDIDATES	126
TECHNOLOGY	4969	ANALOGUE	393	FABRIC	126
PROTOCOL	4742	CHIPS	392	ROTATION	126
SOFTWARE	4575	DESKTOP	391	VAST	126
INTERNET	4531	TRANSPARENT	389	PLUG	124
LAYER	4425	EMAIL	387	REPLICATED	124
INPUT	4347	COEFFICIENT	382	WORKSHOP	124
TRAFFIC	4345	LOOPS	378	INEXPENSIVE	123
MOBILE	4341	COMPETITIVE	374	FLUX	122
OUTPUT	4139	COLLABORATION	373	FUSION	122
CIRCUIT	3932	MATHEMATICAL	372	PROFILING	122
LINK	3905	PORTABLE	372	INTEGRATOR	121
FIG	3702	EFFICIENTLY	371	LTD	121
SERVER	3649	CAREER	369	COMPROMISE	120
DIGITAL	3595	COLUMN	369	PLOTS	120
INTERFACE	3526	COMPUTE	364	REWRITE	119
DEVICES	3430	PROGRAMMABLE	364	LOS	118
CHANNEL	3212	DOWNLOAD	363	STRENGTHS	117
COMPUTER	3201	DATABASES	361	OVERHEADS	117
COMMUNICATION	3159	PROPRIETARY	359	DISKS	116
COMMUNICATIONS	3144	OFFSET	358	GRANULARITY	116
BANDWIDTH	3119	SENSORS	358	INTERIOR	116
CODE	3112	WORLDWIDE	358	PORTABILITY	116
WEB	2978	DESIGNING	357	STRAIN	116
VOLTAGE	2945	SEC	356	TRANSPORTATION	114
NODE	2853	VERSUS	350	DEPENDENCY	114
DEVICE	2849	XML	348	CF	113
OPTICAL	2822	CD	344	ALIGNMENT	112
SIMULATION	2817	ASSEMBLY	340	AUTOMOTIVE	112
ALGORITHM	2799	CORRELATION	339	GEOMETRIC	112
PROCESSING	2770	CRYSTAL	339	PRIMITIVE	111
COMPONENTS	2727	SYNTAX	337	FACULTY	109
FILTER	2659	SEMANTICS	336	MICROPROCESSORS	109
ARCHITECTURE	2581	TRANSACTION	333	SCALED	109
TRANSMISSION	2544	AXIS	332	AUXILIARY	108
MEMORY	2458	KERNEL	332	MULTIPLEXER	108
NODES	2361	SAN	331	TRIVIAL	108
PACKETS	2308	SUBSET	328	ACKNOWLEDGMENTS	107
HARDWARE	2257	ENCODING	327	CONVERGE	107
TECHNOLOGIES	2137	REUSE	327	DOWNLOADING	106
CIRCUITS	2122	FIBRE	325	RETRIEVE	105
CELL	2089	LAUNCHED	324	HEWLETT	105
SWITCH	2075	HTML	323	DEPICTS	105
COMPONENT	2000	GENERATOR	322	COLUMNS	105
LOGIC	1920	MICRO	322	RECTANGLE	104
PROCESSES	1912	UPDATED	321	BIOLOGICAL	104
CONFIGURATION	1885	VECTORS	321	TC	103
DISTRIBUTED	1824	LINKED	320	OPTIMIZING	103
DESIGNED	1822	PROCESSED	320	ALPHA	103
ALGORITHMS	1777	FABRICATION	317	TEMPLATES	102
REMOTE	1770	IC	317	LINKING	102
CABLE	1715	TI	314	PROPAGATE	101
ELECTRONIC	1698	SENIOR	306	EXTENSIBLE	100
ATM	1639	VELOCITY	305	CONCURRENTLY	99
CLIENT	1601	VERTICAL	302	RIGOROUS	98
STORAGE	1591	TRANSACTIONS	301	OVERLAPPING	98
LINEAR	1590	ARRAYS	298	MIRRORS	98
CHANNELS	1569	NOVEL	298	SYMMETRY	97
VIDEO	1559	UPSTREAM	293	PACKARD	97
REFERENCE	1543	INTERCONNECT	291	LOCALIZED	96
SPECTRUM	1499	PLACEMENT	291	HOMOGENEOUS	96
DATABASE	1451	PLOT	291	ALIGNED	95
DISTRIBUTION	1448	SCALING	291	ELEMENTARY	93

SATELLITE	1401	REDUNDANT	289	ALGORITHMIC	93
LINKS	1394	SIMULTANEOUS	287	EMBEDDING	91
PORT	1392	SOPHISTICATED	287	RESULTANT	89
JAVA	1365	DEGRADATION	286	PENALTY	89
WAVELENGTH	1352	MOS	284	DRAWBACK	89
MEDIA	1322	TUTORIAL	284	MEMS	88
SWITCHING	1315	OUTGOING	279	SPREADSHEET	87
AL	1299	NOTATION	278	SQL	86
ANALOG	1264	THEOREM	278	PROTOTYPING	85
MULTI	1260	OPTIMUM	277	AB	85
TRANSFER	1254	COMPACT	274	BURDEN	84
ANTENNA	1242	MESH	273	ADVENT	84
CHIP	1229	MACRO	272	TRANSFERRING	83
COMPUTING	1227	GRAPHICS	271	BATCH	83
SERVERS	1208	RADIUS	271	TEXTBOOK	81
FILTERS	1207	ADDITIONALLY	270	REPETITION	81
TRANSPORT	1195	ITERATION	269	MONOLITHIC	81
INTERFACES	1187	LAB	267	CC	81
EXECUTION	1186	PROGRAMMER	267	MASSIVE	80
CARRIER	1145	INVERSE	261	PRESENTATIONS	79
PC	1142	DEVELOPER	261	PATENT	78
ET	1116	HUGE	257	PROBES	77
TELECOMMUNICATIONS	1105	ARITHMETIC	256	INTERSECTION	77
PROCESSOR	1093	CLUSTERS	256	DIGITALLY	77
ETC	1081	MISSION	256	DELETE	77
COMPUTERS	1076	SEMICONDUCTORS	256	AMBIENT	77
SYNTHESIS	1054	GRAPHICAL	255	UNEXPECTED	76
BROADBAND	1049	CAPTURE	254	SEMINAR	74
IMAGE	1047	PILOT	254	GRAY	74
VENDORS	1004	PROGRAMMERS	254	VIEWPOINT	73
IEEE	1002	BENCHMARK	248	AWARDS	73
CELLULAR	987	QUEUES	248	OPTIONALLY	72
CELLS	957	QUERIES	247	BREAKDOWN	71
TRANSMITTED	957	FEASIBLE	244	TUTORIALS	70
ENTERPRISE	953	CAPACITORS	243	CAPTURES	70
DISCRETE	949	AIRCRAFT	241	CALCULUS	70
ONLINE	937	MAPPED	240	VERSATILE	69
PLATFORM	935	ORTHOGONAL	240	ROTATING	69
HEADER	903	BARRIER	236	PARTITIONS	69
SILICON	897	CIRCUITRY	236	TERMED	68
COM	890	SEGMENTATION	235	NE	68
FREQUENCIES	885	QUANTUM	234	ENGAGED	68
PROCESSORS	884	PARTITION	232	WORKSHOPS	67
FEEDBACK	883	ACCOMPLISHED	230	GAUGE	67
ARRAY	870	DEPICTED	230	DISPERSED	67
MICROWAVE	848	DIAMETER	230	TINY	66
SEMICONDUCTOR	845	ICS	230	PREDETERMINED	66
LAYERS	839	CONJUNCTION	229	GRAPHIC	66
IMPLEMENTATIONS	826	INTRA	229	FOSTER	66
MODULE	824	NAVIGATION	229	UNDERGRADUATE	65
FUNCTIONALITY	819	ITERATIVE	225	MI	65
CLICK	815	PHYSICS	223	SCALAR	64
BUFFER	813	SCRIPT	222	PROMINENT	64
CO	808	SYMMETRIC	214	AWARDED	64
STATIC	808	TRANSFERRED	214	ASSEMBLED	64
THROUGHPUT	806	HEIGHT	213	EXPERIMENTALLY	63
HTTP	801	ACCESSIBLE	213	DEvised	63
SUB	800	BRIDGING	213	COUNTERPARTS	63
WWW	796	MIN	212	SYNTHETIC	62
MATRIX	795	PROBE	210	INTUITIVELY	62
OVERHEAD	792	HETEROGENEOUS	209	GRAINED	62
OPTIMAL	788	CAD	208	DASHED	62
MICROSOFT	782	ENCODED	208	QUANTIFY	61
CACHE	781	NEIGHBORING	208	SONY	60

LATENCY	777	OPTIMIZE	203	ENTITLED	60
SEGMENT	756	SCANNING	201	VIBRATION	59
DC	749	ATOMIC	200	TRANSFORMING	59
TRANSFORM	745	HORIZONTAL	200	REUSED	59
LOGICAL	744	SUPER	200	LATERAL	59
ARCHITECTURES	740	FABRICATED	199	GENERALIZATION	59
DENSITY	737	IMAGING	199	UNDERWAY	58
TRANSMIT	737	CLUSTERING	197	TRAVERSE	58
GHZ	734	FRACTION	197	INCURRED	58
USAGE	727	ROUTINE	197	MATHEMATICALLY	56
EMBEDDED	721	TRANSFERS	196	LITERALLY	56
RF	716	TRANSFORMATIONS	194	ENVISIONED	55
COMPUTATION	707	MOLECULAR	193	ENCODE	55
DESIGNS	706	COUPLE	192	COUNTERPART	55
UPDATE	702	INSTRUMENTATION	192	APPEAL	55
GATEWAY	694	BUDGET	191	SUBTLE	54
SWITCHED	692	ROM	190	STRINGENT	53
IBM	664	DETERMINISTIC	188	MERIT	53
REFERENCES	664	DIAGRAMS	187	VERTICALLY	52
CONNECTIVITY	661	CE	185	SUBSETS	52
VERIFICATION	632	MINIMIZATION	185	PERTINENT	52
MAPPING	631	CA	182	MO	52
MULTIMEDIA	627	CANDIDATE	182	MANAGEABLE	52
OPTIMIZATION	627	DISTRIBUTIONS	182	SEMINARS	51
DIFFERENCES	614	SCAN	180	DEGRADES	51
CODES	598	DEPENDENCE	179	INTERESTINGLY	50
SIMULATIONS	593	DECOMPOSITION	177	INSENSITIVE	50
DESIGNATED	581	SUPERIOR	177	HOUSEHOLD	50
DUAL	574	PROFILES	176	CAST	50
BINARY	573	SEAMLESS	175	AUGMENTED	49
DEDICATED	560	EMISSION	174	PREDECESSOR	48
CODING	556	PROFESSOR	174	INSPIRED	48
SYNCHRONIZATION	556	INTUITIVE	172	STEREO	47
MODULES	554	SUBSYSTEM	170	REUSABLE	47
OUTPUTS	551	COMPOUND	169	ADVANTAGEOUS	47
MOBILITY	548	TECH	169	RENDER	46
CONCURRENT	544	VACUUM	169	APPEALING	46
EXECUTED	541	IR	168	ROTATE	45
LAYOUT	535	VITAL	168	CUSTOMIZE	45
TRANSFORMATION	525	DENSE	167	INTERCONNECTING	44
LABORATORY	524	DISTRIBUTE	167	GRAPHICALLY	43
TEMPORAL	524	INTERCONNECTED	166	CONCISE	43
INCOMING	523	GRADUATE	164	WATTS	42
INC	519	LAYERED	163	BRACKETS	42
THRESHOLD	518	DEFENSE	162	SHRINK	41
GENERIC	517	ACCESSES	161	ALIGN	41
MAGNITUDE	513	EXECUTES	160	GEOMETRICAL	40
COLLABORATIVE	509	FRAGMENT	160	STIMULATING	39
OVERVIEW	508	FUNDAMENTALS	160	REPETITIVE	39
PEAK	500	TUNNELING	160	OPTIMISE	39
STACK	498	MATHEMATICS	158	DISCONTINUOUS	39
DIFFERENTIAL	497	ROBUSTNESS	158	SWAPPING	38
SIMULTANEOUSLY	496	TERMINOLOGY	156	SPARC	38
SEGMENTS	490	SENSING	155	OBSTACLE	38
IMAGES	489	TEMPLATE	154	INCREMENTALLY	38
SPATIAL	489	CORP	153	ASSEMBLING	38
PROFILE	485	MOTOROLA	152	ADJUSTABLE	38
COUPLING	479	MODULAR	152	STRIVE	37
COUNTER	477	DIRECTIONAL	150	GLUE	37
VS	477	AWARD	149	ATTENDEES	37
SETTINGS	475	POP	149	ENCOMPASSING	36
SMART	475	WORKSTATIONS	149	SCANS	35
INTEROPERABILITY	468	CONSORTIUM	148	PIONEERED	35
DISK	463	TRANSFORMED	148	POTENTIALS	34

QUERY	463	WEBSITE	148	SYNTHESIZING	33
DESIGNERS	461	QUANTITATIVE	147	ACCELERATED	33
UPDATES	461	RELATIONAL	147	RESORT	32
CONFIGURATIONS	459	COMPLIANT	146	PROFOUND	32
INSTALLED	455	ARCHITECTURAL	145	POLYTECHNIC	31
COUPLED	450	BARRIERS	144	HORIZONTALLY	31
ROBUST	449	MOUNTED	144	RESEMBLES	30
PLATFORMS	446	CONCRETE	142	RENDERING	30
PROTOTYPE	446	PROPOSITION	142	CHARACTERIZING	30
COMPRESSION	444	SPANNING	142	PREFERABLY	29
INTEGER	444	VIABLE	142	PORTING	29
COMPUTED	443	WEIGHTED	142	MULTILEVEL	29
CONVERGENCE	441	CAPTURED	141	AGGRESSIVELY	29
SUITE	439	NETWORKED	141	NAIVE	27
HYBRID	437	SPAN	140	INTACT	27
AC	431	ENCAPSULATION	139	DASH	26
THERMAL	428	ULTRA	139	IDEALIZED	25
OPTIMIZED	427	CORRELATED	138	DICTATED	23
VERIFY	427	INCREMENTAL	133	FASCINATING	22
METRICS	421	LABORATORIES	133	ABSORB	22
TM	418	OVERLAP	133	CRUDE	20
COMPILER	417	BASICS	132	DEFICIENCY	14
CONFIGURE	417				

Tabla 8.2 Palabras presentes en todas las áreas.

8.1.2.3 Palabras únicas de cada área.

A continuación se muestran los resultados obtenidos tras clasificar las palabras que sólo aparecen en un área de conocimiento. La siguiente tabla indica, para cada área, el volumen total de formas, el valor de frecuencia máxima y el número de palabras registradas en los rangos de frecuencia señalados. En una segunda tabla se expone el porcentaje que cubren las palabras únicas en cada área.

Áreas de conocimiento	Total de palabras únicas por área	Frec. máxima	Palabras frec = 1	Palabras frec = 2	Palabras frec = 3	Palabras frec entre 4 y 8	Palabras frec entre 9 y 20	Palabras frec entre 21 y 65	Palabras frec >65
01 Electrónica	2.456	65	1.436	615	174	225	111	35	-
02 Arq.Comp	729	48	395	180	49	57	40	8	-
03 Telemática	3.377	217	1.963	548	205	448	146	55	12
04 Señal	1.592	107	1.060	237	105	97	75	177	1
05 Materiales	830	51	545	137	44	73	24	7	-
06 Empresa	1.527	218	1.032	250	79	97	47	20	2
07 Sistemas	816	192	503	95	67	98	34	16	3
0801Esp.Señal	2.219	362	1.312	419	-	300	128	49	11
0802Esp.Telmt	898	150	535	-	42	140	120	55	6
03 + 0802	315	858	-	158	78	72	-	-	7
04 + 0801	237	74	-	56	50	66	42	22	1

Tabla 8.3 Volumen de palabras únicas de cada área.

Áreas de conocimiento	Nº total de palabras	Nº total de palabras únicas	Cobertura
01 Electrónica	20.142	2.456	12,8%
02 Arq.Comptad.	11.328	729	6,4%
03 Telemática	25.768	3.377	13,1%
04 Señal	19.330	1.592	8,2%
05 Materiales	8.662	830	9,5%
06 Empresa	16.368	1.527	9,3%
07 Sistemas	13.081	816	6,2%
0801 Esp.Señal	22.078	2.219	10%
0802 Esp.Telemática	21.787	898	4,1%
03 + 0802	47.555	315	0,6%
04 + 0801	41.408	237	0,5%

Tabla 8.4 Cobertura de las palabras únicas.

Como se puede apreciar en las tablas anteriores, se han añadido dos secciones extra: 03+0802 y 04+0801. Éstas se corresponden, respectivamente, con el área de telemática (03) más su especialidad (0802), y con el área de procesamiento de la señal (04) más su especialidad (0801). La razón por la que se introducen tales secciones reside en el proceso de clasificación, donde se observó que determinadas palabras sólo aparecían en dos secciones, y que éstas coincidían con una especialidad y su área de conocimiento principal. Después del recuento, se ha hallado un número considerable de palabras comunes a las dos secciones y que a la vez sólo se dan en ellas. Por lo tanto, se consideran candidatas a términos técnicos, al igual que el resto de palabras que sólo aparecen en un área de conocimiento: *“one of the ways in which terminology can be automatically identified is in terms of its greater tendency to occur only in a limited range of texts”* (Skehan, 1981 in Aston, 1996). En el Apéndice II se encuentran los listados de palabras correspondientes a cada área. A modo de ejemplo, se muestran a continuación, las 50 primeras palabras de frecuencia mayor de cada área.

01 Electrónica	02 Arq.Computadoras	03 Telemática	04 Proces. Señal	05 Materiales
PHOTORESIST	PROCESSMACHINE	NSSA	SVM	FOAMS
FBDD	COONS	SCWP	RFEMF	COPOLYMERS
LMIS	FACTORYSUPERVISOR	NBMA	CLMS	NANOFIBRES
FLIT	PEAKFPGA	GROUplet	ELIN	EDOT
NISI	GCLP	SHWP	FAck	DICNFL
EPCOS	MPLAB	ISLIP	LSW	AMIDE
WLR	FEEDSTOCK	DCL	FPTA	YB
FLITS	DTO	KAWA	BICONICAL	UVILS
OBDD	HALFBAND	EOMPLS	IAMPS	SINTERING
GTLF	PROCESSMACHINES	EVAL	IFIR	FLEXURAL

RHODE	SEMATECH	COCODOC	LZW	PENTOXIDE
SILICIDE	ULOGIC	ACTIVATABLE	BMG	TETRAHEDRAL
JFET	VCCINT	DRMS	DILATION	ZPL
PPV	APG	TRSS	PSPICE	MICROBALLOONS
PROTEO	ELU	MOZAMBIQUE	ROBOLAB	OCTAHEDRAL
EHP	LOTJOB	GROUPLETS	CELLVIEW	FLUOROETHYLENE
MEH	OPTOCOUPLER	MLINK	QIM	BALDRIGE
BTE	VECTORIZATION	TELETRAFFIC	REVERB	HFAC
JSC	JOBUILDER	TELECITIES	APB	SOLUBILITY
EUV	DIAMONDOID	BFS	EXCELERATOR	RECESSES
TRANSIMPEDANCE	ENTERPRISEEXECUTOR	OGSI	SALINITY	NCLA
SBH	VISUALSTATE	ETHERWARE	OLOFSON	SNOM
BBT	CSAS	DSAS	REGULARIZATION	TMHD
CSTEPS	CXT	DCMM	CHAFLANE	AMIDES
PHOTOCONDUCTOR	CYCLONE	WSRF	GAUSSMETER	TMVS
ELECTRONICAST	EDVAC	FLIPPER	PERIDOGRAM	CHLORINE
FBDDS	MSF	CSH	VITATECH	DIKETONATE
LVCM	QBAR	VIRTUALMEDIA	APPLETVIEW	EEEL
PHOTOCURRENT	RIBOSOME	OPP	BERENGER	LASE
PORES	STIMULATOR	SLAPD	HORTON	LIGAND
DIAPHRAGM	COSYN	ITM	LINC	SINTERED
OPTIMISER	HIERARCHIZATION	PMBOK	MILLIGAUSS	TITANATE
THYRISTORS	OUTFILE	QUIPU	TBL	ADDITIVES
VBE	PSFEM	BAZ	DITHER	DEBONDING
ACCEPTORS	QUICKLOGIC	MOUSETRAP	ESTAR	FOXBOX
CLPF	QUICKSTART	ITUNES	HRN	ISOPROPOXIDE
EKC	REPELLER	CBPDUS	LTRPUB	MAGNESIUM
FEMTOSECOND	SQP	LMP	NWP	POLYTYPIC
INNOS	STARTJOB	MDRM	PML	SOLVENTLESS
DATASHEET	APERIODIC	QUASIQUEOTE	UNCOMPRESSION	SUBOXIDE
EDR	CONV	ARBITRATOR	VOUCHER	DMAC
TOWNES	EXTREMITY	TMC	SWITCHGEARS	FERROCENE
VTH	MOSTEK	CCNAS	AMX	FLAHERTY
WEHP	NSFEM	CITEL	RADIANCES	POLYCONDENSATION
BLOCH	SEQUENCER	CBPDU	CIW	PYRIDINE
KNOWLEDGEBASE	UNF	CSHRC	EROOM	RECOIL
MICROGRAPH	VHTN	ERIS	MULTIPROCESS	COERCIVITY
MONOLAYERS	DELAYOR	NSSAS	MUMETAL	COPOLY
SECOP	DFE	REBOL	OERSTED	DIBLOCK

Tabla 8.5a Únicas del área.

06 Org. Empresas	07 Sistemas	03 + 0802	04 + 0801
HARASSMENT	DMCS	LSAS	LNA
GAM	LRE	LSA	INTERFERERS
ROAMABOUT	NPID	PMI	LTR
GLOBALIZING	MDC	SNA	PFD
HABITUS	DIOS	APPLETALK	AOC
TEMPORALITY	SCALEA	IGRP	MMSE
KTNET	LAPI	ELITECONNECT	AWR
IDEALISM	TLG	ANSITEST	HFSS
RELATIVISM	UMM	ASPIRE	RADIOMETER
INTERNATIONALS	TRIMPI	BAKEOFFS	AZIMUTH
SOLSOFT	AIC	BONCHEK	ACA
URIS	HCML	BOOMERS	FEEDERS
HOFSTEDE	SCML	BROWSED	SNRS
HARASSERS	LEBEP	CIRAVEGNA	TW
HARASSING	LQR	COLLABORATIONAWARE	METEOROLOGICAL
MADSEN	SIMADA	CONSING	PMR
DAIRY	DRG	COQ	PROTECTORS
FRINK	HOMEOMORPHIC	COURTRIGHT	MISADJUSTMENT
PACKETSHAPER	CDC	CRONAN	WSS
BYSTANDER	OCCAM	DEFAULTDESTINATION	RADIONAVIGATION
IMOBILE	DIVERGENCES	DENBIGH	BACKSCATTER

BUSHBY	MEIKO	DESYNCHRONIZATION	DIURNAL
CATALOGEXAMPLE	RIPPER	DIFFUSIONISM	SMARTPHONE
CONSULTCO	MSDF	DIRECTFETS	AMSU
FIELDNOTES	ASMOD	DISSENTING	BOS
HARASSER	DUES	DRAWABLEOBJECT	ONU
SALCOMP	MBPC	DRAWHOUSE	POSTPROCESSING
CODEVELOPMENT	TRIMPIS	DXI	WELDED
XEC	CIRCAL	EGAN	QP
BIGSEARCH	CWB	EQUALSIGNORECASE	UNDERSAMPLING
POSITIVIST	ACCEPTANCES	FNF	LATITUDES
BLOGS	DEMONIC	GIGAFAST	LNAS
INTERPRETIVIST	FAULTYSYSTEM	HARINK	VOLTERRA
SURESYN	GDM	INSTANTOFFICE	FLANGE
BOWES	PARMACS	INTS	MICROCELLULAR
GLOBALIZE	QSM	LATECOMING	NOAA
PACKETEER	REDHAWK	LIQUIDATED	POLARIMETRIC
SMOPEC	DUNEDIN	LOCATEREGISTRY	MICROCELL
TETLOCK	HOMEOMORPHISM	MESSAGESTREAMS	MICROSTRIPS
BOLAND	IPARS	MOLDING	NDA
BPR	NEUROFUZZY	NOAM	SONNET
POSITIVISM	PACMOD	NODOSE	AMATEURS
QCS	VSPACE	NONCOMMERCIAL	DOA
MVQ	SCALEA'S	NPIMAGE	DPSK
EPISTEMOLOGICAL	WONDERWARE	NRK	PDC
EPQ	MULTIPROGRAMMED	NSB	PICONET
GENERALIZABILITY	THREAD'S	OTALK	CYCLOSTATIONARY
GLOBALIZED	COMPRESSIONS	PALLADIUM	DECORRELATING
LMX	DOWDEN	PERMUTABLE	EMFS

Tabla 8.5b Únicas del área.

El tipo de información que se ha conseguido aportar sobre el léxico de las telecomunicaciones en este estadio es de diversa índole. En primer lugar, gracias a las listas de corte y una laboriosa limpieza manual, se ha conseguido reducir considerablemente el tamaño inicial de formas. Esto significa que entre el vocabulario especializado, además del vocabulario académico y general más frecuente, se intercala una gran cantidad de palabras no válidas que dificultan el reconocimiento del vocabulario específico.

Tras tamizar el listado de frecuencia inicial, aún resta un importante volumen de palabras pertenecientes a los tres niveles léxicos. No obstante, se dispone del valor de la frecuencia absoluta de las palabras y su distribución por áreas de conocimiento. Asimismo, se han identificado las palabras que se presentan en todas las áreas y aquellas que sólo aparecen en una sección.

Sin embargo, estos datos no revelan resultados esclarecedores. Entre las palabras de alta frecuencia y alta distribución se encuentran tanto palabras técnicas y semitécnicas, como vocabulario general relacionado con las telecomunicaciones o con otras temáticas. El mismo fenómeno se reproduce entre las palabras de baja o alta frecuencia que pertenecen sólo a un área. A partir de este momento, el siguiente paso implicaría el estudio cualitativo de cada palabra en su contexto para descifrar el nivel léxico al que pertenece, puesto que no se dispone de otros datos que desvelen tal condición. En los

cuadros siguientes se encuentran algunos ejemplos de las palabras que responden a los diferentes tipos de comportamiento y nivel léxico, en función de los valores de frecuencia absoluta y distribución. La selección se ha efectuado cualitativamente y al azar:

Distribución alta	Palabras técnicas	Palabras semitécnicas	Palabras generales
Frecuencia alta	<i>ATM, bandwidth, latency, AL, java, wavelength, DC</i>	<i>software, switched, chip, network, array, interface, mapping</i>	<i>Communication, storage, client, peak, image, system, link, technology</i>
Frecuencia baja	<i>MO, gray, scalar, Mi, NE, calculus</i>	<i>Assembling, watts, Customize</i>	<i>Deficiency, obstacle</i>

Cuadro 8.1 Clasificación cualitativa en función de la frecuencia y distribución.

Distribución baja	Palabras técnicas	Palabras semitécnicas	Palabras generales
Frecuencia alta	<i>Etherware, coons, JSC, nanomachinery, grouplet, DIOS</i>	<i>Salinity, sacks, enforcer, interferes, feeders, teletraffic</i>	<i>Pores, extractions, bumpers, foams, payout</i>
Frecuencia baja	<i>ACMOS, radioscope, alkylidithol, outband</i>	<i>Saline, teleoperators, abortion, antiviral, aerospatial</i>	<i>Basements, wrinkle, affirming, weary, monument, addressees</i>

Cuadro 8.2 Clasificación cualitativa en función de la frecuencia y distribución.

Los valores de frecuencia absoluta y distribución, aunque no son suficientes para identificar los diferentes tipos de vocabulario con claridad y eficiencia, demuestran ser criterios básicos acertados para tener en cuenta a la hora de decidir el contenido léxico que debe formar parte de un curso de inglés especializado. Especialmente, serían de gran utilidad aquellas palabras de contenido de alta frecuencia, que a la vez están presentes en todas las áreas (valor de distribución máximo), ya que formarían un conjunto de palabras común y habitual en el lenguaje de las telecomunicaciones. Por otro lado, las palabras de alta incidencia que sólo aparecen en un área reflejan pautas características de términos técnicos.

No obstante, para mejorar los resultados y optar a una clasificación del vocabulario de forma automática, es conveniente comparar el comportamiento estadístico de las palabras en el corpus técnico con su comportamiento en un corpus general y aplicar métodos más sofisticados.

8.2 Clasificación del vocabulario mediante comparación de corpus.

De la literatura consultada, sólo Teresa Mihwa Chung (2003) describe el procedimiento seguido con la suficiente claridad para que se pueda replicar. Chung identifica los términos técnicos del lenguaje de anatomía, comparando un corpus

específico con un corpus general. En su estudio, comparte la clasificación cualitativa que plantea Nation (2000 y 2003) sobre las palabras del discurso; define término como una palabra cuyo significado es específico del dominio donde se utiliza; y toma como unidad de análisis el tipo de palabra o forma. Para identificar los términos cuantitativamente, utiliza como herramienta un cociente o ratio, partiendo de dos supuestos:

1. La frecuencia de las palabras no técnicas (*non-term*) debería ser aproximadamente la misma en el corpus específico y en el general.
2. La frecuencia de los términos técnicos debería ser mucho más alta en el corpus específico que en el general.

En el presente estudio se replica el método propuesto por Chung para identificar los términos. El procedimiento se compone de los siguientes pasos:

1. Obtención de los listados de frecuencia de cada corpus.
2. Eliminación de palabras del corpus general que no aparecen en el específico.
3. Normalización de las frecuencias de las palabras para hacerlas comparables y calcular el ratio o cociente de las frecuencias. De este modo, se obtiene un valor que indica cuánto mayor es la frecuencia de una palabra en el corpus técnico que en el general.
4. Agrupación de las palabras del corpus específico que comparten características similares en términos de frecuencia y ratio.
5. Revisión de los grupos obtenidos y evaluación de las palabras en referencia a los criterios cualitativos que las definen como términos o no términos.

8.2.1 Cálculo de la frecuencia relativa y el ratio.

Los listados de frecuencia se obtienen con el programa WordSmith y después se transfieren a una hoja de Excel para realizar los cálculos requeridos y eliminar las palabras que no aparecen en el corpus técnico. Chung conserva las palabras comunes a los dos corpus, más las únicas del especializado, aplicando el concepto de distribución de las palabras en este sentido. Define distribución como “*a count of the number of different corpora, sub-corpora or texts in which the same word form occurs.*” Al decantarse por la primera opción, no otorga a la distribución de las palabras un valor numérico que participe en las operaciones destinadas a identificar términos.

Como el tamaño de los corpus es diferente, es preciso normalizar las frecuencias para poder comparar los valores, operando así con la frecuencia relativa y no con la frecuencia absoluta. La frecuencia normalizada se obtiene tomando el valor de la frecuencia de cada palabra, dividiéndola entre el número total de palabras del corpus al

que pertenece y multiplicándola por 100. Si se toma como ejemplo la palabra *circuit*, su frecuencia normalizada en cada corpus es la siguiente:

$$\text{Corpus técnico: } 3.931 \div 5.533.705 \times 100 = 0,071037397$$

$$\text{Corpus general: } 520 \div 21.016.312 \times 100 = 0,002474269$$

A continuación, se divide el valor del corpus técnico entre el valor del corpus general, para saber cuántas veces la palabra es más o menos frecuente en el corpus técnico que en el general. Siguiendo con el ejemplo de *circuit*:

$$\text{R: Ratio: } 0,071037397 \div 0,002474269 = 28,71046352$$

El resultado muestra que *circuit* es 28,71046352 veces más frecuente en el corpus técnico que en el general.

Para interpretar el valor del ratio, Chung indica los siguientes parámetros:

- Ratio = 1 significa que dos palabras tienen la misma frecuencia.
- Ratio >1 significa que la palabra es más frecuente en el corpus específico que en el general.
- Ratio <1 significa que la palabra es menos frecuente en el corpus específico que en el general.
- Ratio = infinito significa que la palabra no aparece en el corpus general.

En el Apéndice III se encuentra el documento de Excel con el listado de palabras compartidas y el total de los valores obtenidos: frecuencia absoluta, frecuencia relativa y ratio.

8.2.2 Agrupación de las palabras con características similares.

El corpus general contiene 176.237 formas o tipos diferentes de palabras y el corpus de teleco 59.826; ambos comparten 34.395 formas y 25.431 sólo aparecen en el técnico. Con el fin de identificar el vocabulario específico característico de las telecomunicaciones, se someten a análisis sólo las palabras presentes en el corpus especializado, desechando aquellas que aparecen únicamente en el general.

A la luz de los resultados conseguidos en la investigación realizada por Chung, las palabras se han ordenado en seis grupos según su comportamiento estadístico, en función de la frecuencia absoluta y el valor del ratio.

- Grupo 1: Palabras que sólo aparecen en el específico y ocurren sólo una vez.
- Grupo 2: Palabras que sólo aparecen en el específico y su frecuencia es ≥ 2 .
- Grupo 3: Palabras que aparecen en los dos corpus, su frecuencia en el corpus técnico es ≥ 12 y el ratio es ≥ 50 .
- Grupo 4: Palabras que aparecen en los dos corpus, con frecuencia 1 en el específico y ratio < 50 .
- Grupo 5: Palabras que aparecen en los dos corpus, con frecuencia baja en el específico y ratio < 50 .
- Grupo 6: Palabras que aparecen en los dos corpus, con frecuencia moderada o alta en el específico y ratio < 50 .

Tras evaluar las palabras contenidas en cada apartado y contrastarlas con la clasificación cualitativa *Rating scale approach*, Chung concluye que los grupos 1, 2 y 3 se identifican como términos y los grupos 4, 5 y 6 como no términos. Las palabras catalogadas estadísticamente en los grupos 1, 3, 5 y 6 coinciden en un 77,1%, 88,6%, 85,7% y 77,1% respectivamente, con la clasificación cualitativa previa realizada por el experto en la materia. Los resultados obtenidos en los grupos 2 y 4 son incluso mejores, 91,4% y 94,3% respectivamente. Con esta información se obtienen las siguientes conclusiones:

- a) Para distinguir los términos por medio de una comparación estadística es conveniente que la frecuencia mínima de las palabras en el corpus técnico sea 2.
- b) El criterio frecuencia > 2 y ratio infinito indica de forma bastante fiable (91,4%) que la palabra es un término técnico.
- c) El ratio 50 como valor de corte es un criterio apropiado para distinguir las palabras que son términos técnicos de aquellas que no lo son.

El método de clasificación cualitativa *Rating scale approach* se sirve de las características formales y semánticas de las palabras, y las distribuye en cuatro niveles, dependiendo del grado de especificidad del significado de la palabra en el dominio específico. Las palabras clasificadas en los niveles 3 y 4 se consideran términos.

- Nivel 1: Palabras cuyo significado no está relacionado con el dominio específico; su significado es independiente de la materia en la que aparecen.
- Nivel 2: Palabras cuyo significado está mínimamente relacionado con el dominio específico.
- Nivel 3: Palabras cuyo significado está estrechamente relacionado con el dominio específico. Estas palabras pueden ser términos en un dominio

específico y aparecer igualmente en la lengua general o en otros campos del saber, con el mismo significado pero sin ser términos en tales campos.

- Nivel 4: Palabras cuyo significado y uso está restringido a un dominio específico, y no son conocidas en el lenguaje general.

No obstante, se debe destacar que la última decisión respecto a la inclusión de cualquier palabra en una categoría se toma por medio de criterios cualitativos, bien sea consultando a un experto en la materia o bien estudiando los contextos en los que se encuentre dicha palabra. El método que Chung propone es un intento de obtener una clasificación del vocabulario válida, mediante un procedimiento cuantitativo y automático, con un nivel de fiabilidad razonablemente alto.

8.2.2.1 Grupo 1: Palabras que sólo aparecen en el específico y ocurren sólo una vez.

Debido al elevado número de palabras que obtienen ratio infinito (palabras que sólo se encuentran en telecomunicaciones), se han dividido en dos grupos dependiendo de su frecuencia absoluta. De las 25.431 formas que sólo aparecen en el corpus de telecomunicaciones, 13.551 presentan frecuencia 1.

Una vez que se han obtenido los grupos, es necesario realizar una limpieza para eliminar las palabras que no son válidas, tales como nombres propios, apellidos, palabras en otros idiomas, erratas, etc. Este filtrado afecta en mayor grado al grupo 1, el cual se reduce a menos de la mitad, quedando en un tamaño final de 4.985 formas (Apéndice III: Lista filtrada de las palabras del grupo 1).

Chung sostiene que las palabras que aparecen una sola vez en el corpus específico y no se presentan en el corpus general, afectan al grado de fiabilidad en la comparación estadística. Si se incluyen como términos, el porcentaje de coincidencia con el método de clasificación cualitativa se reduce ligeramente. Por el contrario, si se excluyen, el porcentaje de fiabilidad aumenta pero se pierden términos. La autora opta por excluirlos en favor de la funcionalidad y precisión en los resultados. Normalmente, se suele prescindir de los *hapax legomena* en este tipo de estudios (Scout, 1997; Vangehuchten, 2000; Curado, 2001), bajo el alegato de que se necesita más de una incidencia para poder determinar un tipo de comportamiento.

8.2.2.2 Grupo 2: Palabras que sólo aparecen en el específico y con frecuencia ≥ 2 .

El segundo grupo reúne las palabras que sólo aparecen en el corpus de teleco y que presentan una frecuencia mayor o igual a 2, siendo *OSPF* la forma más incidente (Frec: 1.283). Estos criterios indican de forma bastante fiable (91,4%) que la palabra es un término técnico.

El volumen de palabras asciende a 11.880, que tras el correspondiente filtrado se reduce a 8.118 (Apéndice III: Lista filtrada del grupo 2). La mayoría de las palabras eliminadas en este grupo han sido aquellas que llevan un apóstrofe señal del genitivo sajón. Suelen ser palabras generales como *world's*, *student's*, *week's*, etc., o palabras que se encuentran en el mismo grupo pero sin apóstrofe, por ejemplo: *FPGA's*, *VHDL's*, *foundry's*, etc.

8.2.2.3 Grupo 3: Palabras que aparecen en los dos corpus, su frecuencia en el corpus técnico es ≥ 12 y el ratio es ≥ 50 .

Las palabras que obtienen un ratio de 50 coinciden en tener una frecuencia mayor o igual a 12. El volumen inicial de este grupo es de 806 palabras, de las cuales sólo 18 resultan ser no válidas, quedando en un tamaño final de 786 formas (Apéndice III: Lista filtrada del grupo 3). El porcentaje de fiabilidad alcanzado es de un 88,6%, ligeramente inferior al del grupo anterior. *IP* es la forma más incidente (Frec: 5.240 y R: 995) y *LSAS* la forma que obtiene el valor más alto de ratio (Frec: 858 y R: 3.258); *LSAS* es 3.258 veces más frecuente en el corpus técnico que en el general.

En este grupo se encuentran palabras catalogadas estadísticamente como términos, pero que cualitativamente no lo son, por ejemplo: *optionally*, *initialize*, *subspace*, *multivariable*, *robustness*, etc. Aunque en este apartado no se realice un análisis detallado del comportamiento de tales palabras, este hecho merece una breve explicación. La razón principal por la que se seleccionan como vocabulario especializado es de carácter cuantitativo, es decir, la frecuencia en la que se presentan en el corpus técnico es lo suficientemente alta en comparación con su frecuencia en el corpus general: *optionally* (Frec. técnica:72, Frec. general:4, Ratio:68), *initialize* (F.t:68, F.g:2, R:129) y *robustness* (F.t:158, F.g:12, R:50). Después, al examinar sus contextos, se advierte que algunas de estas palabras forman parte de términos compuestos como *subspace methods* y *multivariable systems*.

8.2.2.4 Grupo 4, grupo 5, grupo 6 y grupo 7.

Las palabras restantes se han dividido en diferentes grupos de acuerdo con un comportamiento estadístico determinado, tal y como se ha expuesto anteriormente. Sin embargo, todas ellas obtienen un valor de ratio inferior al preestablecido ($R < 50$) para que se les pueda considerar como miembros de la terminología. Incluso 18.437 formas, independientemente de su frecuencia absoluta, son menos frecuentes en el corpus técnico que en el general. Éstas se han registrado en un grupo nuevo (Grupo 7), en vez de excluirlas directamente. Se trata de palabras cuyo significado no está relacionado con el dominio específico y que se pueden encontrar en cualquier tipo de texto (Apéndice III: Lista del grupo 7).

El grupo 4 contiene 2.338 palabras de frecuencia 1 en el corpus técnico, oscilando su frecuencia en el corpus general entre 1 y 3, al igual que el valor del ratio (Apéndice III: Lista del grupo 4). Estas palabras no son relevantes debido a su escasa incidencia en los dos corpus; la mayoría son nombres propios, palabras generales y de otros idiomas. Respecto a las palabras del grupo 5, a diferencia del grupo anterior, presentan una frecuencia mayor que 1 en el corpus técnico, pero aún demasiado baja. Los valores del ratio no alcanzan la cifra de corte, aunque algunas formas consiguen niveles considerablemente altos: 41,77 y 45,57. Éstas revelan una frecuencia de 11 y 12 respectivamente en el corpus técnico y comparten frecuencia 1 en el corpus general.

A ninguna de las palabras incluidas en los grupos 4, 5, 6 y 7 se le atribuye la categoría de término especializado. Sin embargo, es conveniente revisar las palabras que se encuentran en el límite inmediatamente inferior al ratio de corte, dado que dicho ratio es aproximado. Tales palabras se localizan en el grupo 6 (frecuencia moderada o alta en el corpus específico y ratio <50: Apéndice III). El valor máximo de ratio hallado es 49,80 (*encryption*), el cual va descendiendo paulatinamente hasta alcanzar el valor mínimo = 1,000067023 (*degrees*). Cuando el ratio de una palabra es = 1, quiere decir que la frecuencia en los dos corpus es la misma. Esto sucede con 1.996 palabras, cuyo ratio oscila entre 1,000067023 y 1,99888083, y que suelen ser funcionales o generales, por ejemplo: *without, of, dictionary, new, cannot, empty, routine, local, boxes, etc...*

Las diferencias entre las frecuencias se van haciendo cada vez más notables a medida que el valor de ratio aumenta, observándose a la vez una creciente relación de las palabras con el dominio especializado. Éstas se corresponden con el nivel 2 de la *Rating scale approach*, es decir, palabras cuyo significado está mínimamente relacionado con el dominio específico, y que por su elevado índice de presencia son imprescindibles. Algunos ejemplos son: *program* (R=4), *website* (R=5), *oscillation* (R=7), *calculus* (R=8), *appliance* (R=9), *cables* (R=10), *etching* (R=11), *wired* (R=12), *microsystem* (R=13), *oscilloscope* (R=14), *laser* (R=15), *microwave* (R=16), *implementors* (R=17), *encapsulated* (R=18), *internet* (R=19), *carrier* (R=20), *telecommunications* (R=21), *transmissions* (R=22), *milliseconds* (R=23), *downloads* (R=24), *optics* (R=25), *hyperlinks* (R=26), *modulator* (R=27), *parabolic* (R=28), *synthesizer* (R=29), *layer* (R=30), *hacker* (R=31), *satellites* (R=32), *diffusion* (R=33), *unidireccional* (R=34), *buffers* (R=35), *bipolar* (R=36), *network* (R=37), *electron* (R=38), *browsers* (R=39), *arrays* (R=40), *nanotechnology* (R=41), *URL* (R=42), *WIFI* (R=43), *wavelets* (R=44), *dielectric* (R=45), *concurrent* (R=46), *micron* (R=47), *filter* (R=48), etc...

Los valores que más se aproximan al ratio de corte oscilan entre 49,372 (*simulating*) y 49,806 (*encryption*), abarcando un total de 35 palabras. Entre ellas se aprecian varios casos. En primer lugar se encuentran las palabras que cumplen con las características

generales del grupo, como *equipments*, *computable*, *deactivation*, *histograms*, etc. Por otro lado, se observan tres tipos de particularidades: palabras con frecuencia 1 en el corpus general, terminología y terminología generalizada.

8.2.2.5 Palabras con frecuencia 1 en el corpus general.

Las siguientes siglas se encuentran entre las palabras de frecuencia 1: *DCF*, *DFE*, *IEC*, *PEL*, *QEF*, *RTU* y *SPX*. En estos casos es necesario examinar los contextos en los que aparecen, tanto en el corpus técnico como en el general. Si el origen de las siglas coincide en los dos corpus, permanecerían en este grupo. De lo contrario, si se refieren a conceptos distintos, podrían pasar a ser términos técnicos puesto que su ratio sería infinito. Esto sucede con *DCF* (*Distributed Coordination Function*), *DFE* (*Decision feedback equalizer*), *QEF* (*Quasi-error free*), *RTU* (*rooftop unit*) y *SPX* (*Sequenced packet exchange*). A continuación se muestra, a modo de ejemplo, los contextos que permiten identificar a *DCF* como vocabulario especializado:

Corpus Lacell:

No games. Glen Burnie. 38253 Pretty, blonde DCF, 39, 5'6", attractive, long blonde hair. Seeking...

Corpus Teleco:

AC protocol is called the Distributed Coordination Function (DCF) which uses slotted CSMA with RTS/CTS for DCF (Distributed Coordination Function) of IEEE 802.11. DCF is based on a Carrier Sense Multiple Access with Before W-CHAMB is introduced, we discuss the DCF (Distributed Coordination Function) of IEEE 802.11...

Al comprobar el contexto en el que aparecen palabras como *multiplexors* y *oligomers*, se descubre que son términos especializados. Probablemente, se hallan también en el corpus general, como parte de las muestras de textos técnicos introducidas con el fin de conseguir la representatividad temática de la lengua inglesa.

En general, es conveniente revisar todas las palabras de frecuencia alta o moderada en el corpus técnico (frec.>12), que presentan frecuencia 1 en el corpus general y cuyo ratio es <50.

8.2.2.6 Terminología.

Encryption (49,806) es la palabra que se encuentra justo en la frontera que separa el vocabulario especializado del no especializado, según criterios estadísticos cuantitativos. Sin embargo, cualitativamente, *encryption* se puede considerar término especializado. Éste no es el único caso, incluso en niveles de ratio inferiores, es posible localizar otras palabras especializadas como *DSPs*, *MOFET/S*, *wavelets*, *PID*, *transceiver*, etc. Una vez más, es conveniente recordar que el ratio de corte es un valor aproximado, al igual que el límite que intenta aislar las lenguas de especialidad y la lengua común.

8.2.2.7 Terminología generalizada

Entre los valores más altos de ratio se encuentran palabras como *circuitry*, *HW* (*hardware*) y *sms*, e incluso entre valores menos elevados aparecen *pdf* (R=47), *microwaves* (R=44), *mousetrap* (R=41), *nanotechnology* (R=41), *megabit/s* (R=38/40), *transmitter* (R=24), *websites* (R=12), *software* (R=12), etc., todas ellas reconocibles como propias de las telecomunicaciones, pero que no se han clasificado como especializadas en términos estadísticos. Cabré (1993:167) ofrece una explicación a este fenómeno argumentando que la gran divulgación de las materias especializadas, gracias a los medios de comunicación y la enseñanza, ha hecho que la terminología asociada forme parte de la terminología general que conoce hoy en día un hablante medianamente informado.

8.2.3 Conclusión.

Como se infiere de los apartados anteriores, el presente análisis se ha desarrollado a la luz de los hallazgos conseguidos por Chung (2003) en la aplicación de un método que facilite la distinción del vocabulario especializado, de manera automática y con un porcentaje alto de fiabilidad (86%). Los resultados obtenidos indican que los términos técnicos de la ingeniería de telecomunicación se han registrado principalmente en los grupos 2 y 3, con la opción de estimar las palabras del grupo 1, dependiendo de la finalidad que se persiga. El total de términos técnicos asciende a 8.904 formas, más 5.043 opcionales (Cuadro 8.3 y Figura 8.1). Estas cifras se refieren al conjunto del vocabulario específico en términos de Nation (2001) y Chung (2003), es decir, al vocabulario técnico con sus diferentes grados de especialización.

Grupo	Tamaño inicial	Frec.	Ratio	Formas no válidas	Nº final de formas
1	13.551	= 1	∞	8.508	5.043
2	11.880	≥ 2	∞	3.762	8.118
3	806	≥ 12	≥ 50	20	786
4	2.338	= 1	<50	-	=
5	5.655	<12	<50	-	=
6	7.071	≥ 13	<50	-	=
7	18.437	x	<1	-	=
				Total formas técnicas: 13.947	

Cuadro 8.3 Resumen del volumen de palabras obtenido en cada grupo.

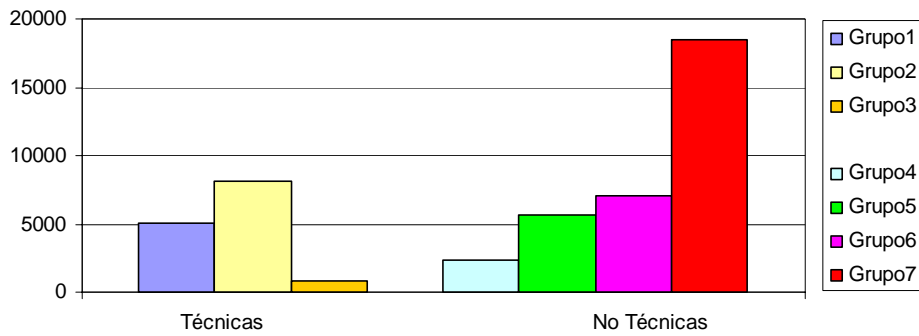


Figura 8.1 Volumen de palabras obtenido en cada grupo.

La selección realizada cuantitativamente comprende tanto palabras técnicas que pueden aparecer en usos no especializados del lenguaje general y palabras generales que adquieren un significado específico, como términos técnicos únicos del dominio específico. Todos ellos definen el perfil léxico del lenguaje de especialidad: “*The lexical level of special languages must be defined as including all items which have special reference, regardless of whether they have general reference or not*” (Sager, 1980:230).

Si se pretenden alcanzar niveles más altos de fiabilidad en la clasificación de términos, es necesario que un experto en la materia concerniente revise cualitativamente los resultados conseguidos. No obstante, tal y como afirma Cabré (1993:168):

Intentar establecer una frontera nítida entre las lenguas de especialidad y la lengua común es una tarea imposible, como lo es querer delimitar de forma rígida las fronteras entre las distintas especialidades, o pretender asignar un término, por sistema, a una sola temática.

Así pues, resulta aceptable admitir una clasificación aproximada, que distingue léxico general de términos especializados, mediante criterios estadísticos y con una alta tasa de fiabilidad.

8.3 *Keywords* o palabras clave del corpus.

En esta etapa de la investigación, se procede a analizar el léxico de las telecomunicaciones en función del grado de relevancia de las palabras que integran dicho sublenguaje, con el fin de obtener un trazo más que permita esbozar su perfil léxico.

El grado de relevancia o *keyness* se obtiene por medio de la aplicación de un test estadístico denominado técnica de razón de verosimilitudes (*Log Likelihood test*), disponible dentro del paquete de utilidades de WordSmith: KeyWords Tool. Esta herramienta compara la frecuencia de las palabras de un texto con un corpus de referencia y genera listados de *keywords* o palabras clave. Scott define *keyword* como

“a word which occurs with unusual frequency in a given text” (Scott, 1997). Las palabras clave son aquellas cuya frecuencia es inusualmente alta o baja en comparación con una determinada norma. El corpus general de mayor tamaño establece la norma con la cual se contrasta el corpus técnico.

Conocer los términos técnicos propios de la especialidad es fundamental para conseguir una comunicación efectiva, principalmente en las situaciones más específicas: “El carácter monoreferencial de los términos desempeña un papel clave en la precisión y univocidad de la comunicación especializada” (Cabré, 1993:167). En realidad, si se desconoce la terminología, no se llega a comprender la materia en sí. Sin embargo, el hecho de que una palabra sea término específico y su uso esté restringido al dominio al que pertenece, no significa que sea representativa. Resulta sencillo ilustrar este argumento, si se toman como ejemplo palabras de baja frecuencia como *fastchip* (frec.7) o *bimos*, *polyspectra*, *securID* y *thinkpad* (frec.4); o incluso palabras como *bootable*, *vectorizable* y *axially* que aparecen sólo una vez y cada una en un área distinta. Todas estas palabras son términos estrechamente relacionados con las telecomunicaciones, pero su incidencia no es comparable con *wireless*, *routers*, *voip* o *wan* cuyas frecuencias son 4.082, 1892, 452 y 580 respectivamente. Estos últimos ejemplos, considerados cualitativamente términos técnicos (*voip* y *wan*) o semitécnicos (*wireless* y *routers*), ejercen una gran incidencia y su uso está más extendido. Por lo tanto, sería conveniente y útil averiguar qué palabras es más probable encontrar en este ámbito o cuáles son más necesarias para conseguir una comunicación efectiva.

8.3.1 Listados de *Keywords*.

El programa WordSmith dispone de dos tests estadísticos para calcular las palabras clave: *Chi-square* y *Log Likelihood test*. Sin embargo, de acuerdo con Dunning (1993) y Scott (1998) el segundo test sería más adecuado en este caso: “*Log Likelihood test, gives a better estimate of keyness, especially when contrasting long texts or a whole genre against your reference corpus.*” Por lo tanto, teniendo en cuenta las características de las muestras que se van a analizar, se ha optado por la aplicación de la técnica de razón de verosimilitudes (*Log Likelihood test*). Asimismo, este procedimiento es ampliamente aceptado y concierne con numerosos estudios realizados (Aston, 1996; Hoey, 1997; Bower, 1998; Tribble, 2000; Curado, 2001; Stubbs, 2001; Jirapa, 2005; etc...) donde se prueba su validez: “*By developing the Keyword program [...], Scott has provided an adequate and robust means for identifying statistically prominent words in a text or collection of texts*” (Tribble, 2000:4).

Para identificar las palabras clave, se toma el listado de frecuencia del corpus de telecomunicaciones más el listado de frecuencia del corpus general, y se transportan a la herramienta *Keywords* en WordSmith. Al aplicar la técnica de razón de verosimilitudes

se detecta si la frecuencia de una palabra en el corpus técnico es significativamente mayor o menor que la frecuencia de la misma en el corpus general. Después, las palabras identificadas como clave, se presentan ordenadas de mayor a menor en función del grado de relevancia o *keyness*.

El programa registra un total de 16.000 palabras clave, de la cuales 12.602 son positivas y 3.398 negativas. Las palabras clave positivas son aquellas cuya frecuencia es significativamente más alta en el corpus técnico que en el general. El valor máximo de *keyness* asociado a una palabra es de 41.784,6 (*network*) y el mínimo es de 10,8 (*broad*). Igualmente, cada valor se ha obtenido con un margen de error que oscila desde 0 a 0,000997, consiguiendo un total de 7.815 palabras clave positivas con margen de error igual a 0. Respecto a las palabras clave negativas, su incidencia es significativamente menor en el corpus específico que en el general. El valor máximo negativo de *keyness* asociado a una palabra es de -10,9 (*educating*) y el mínimo es de -29,8 (*necklace*). El margen de error correspondiente al valor máximo negativo de *keyness* es de 0,000982, el cual va descendiendo paralelamente al valor de relevancia hasta alcanzar un margen de error 0.

El resultado del test responde a la pregunta de investigación planteada, reconociendo aquellas palabras que, estadísticamente, se utilizan en el inglés de las telecomunicaciones con mayor frecuencia que en inglés general, independientemente del nivel léxico al que pertenezcan. Las palabras clave positivas aportan una buena indicación del contenido del texto, en términos de Scott (1998): “*positive keywords give a good indication of the text's aboutness*”. En el Apéndice IV se halla la lista completa de palabras clave. No obstante, debido al elevado número de palabras, se muestran a continuación las primeras 150 palabras clave positivas:

N	Word	Telc. Freq.	Lacell Freq.	Keyness Log L.	N	Word	Telc. Freq.	Lacell Freq.	Keyness Log L.
1	NETWORK	16.649	1.686	41.784,60	76	EACH	10.224	14.172	5.519,00
2	DATA	14.613	2.787	31.852,20	77	COMMUNICATION	3.159	1.130	5.490,80
3	SIGNAL	7.022	641	17.922,60	78	TCP	1.717	12	5.248,00
4	SYSTEMS	9.479	3.000	17.377,70	79	USE	10.255	14.866	5.144,70
5	IP	5.239	20	16.182,10	80	CONFIGURATION	1.885	106	5.134,10
6	NETWORKS	5.832	463	15.204,90	81	CODE	3.112	1.249	5.121,80
7	SYSTEM	12.624	8.707	14.831,60	82	ADDRESS	3.951	2.416	5.070,10
8	USER	6.292	903	14.725,80	83	FUNCTION	3.380	1.677	4.960,60
9	PROTOCOL	4.742	139	13.677,70	84	HARDWARE	2.257	423	4.939,90
10	DESIGN	7.701	3.313	12.237,80	85	TECHNOLOGIES	2.137	329	4.919,50
11	APPLICATIONS	5.414	934	12.117,70	86	ALGORITHMS	1.777	102	4.828,80
12	ROUTER	3.910	25	11.974,40	87	ATM	1.639	35	4.817,30
13	USING	9.214	5.576	11.914,70	88	TYPE	4.612	3.750	4.716,70
14	IS	94.630	234.895	11.589,10	89	PROVIDES	3.257	1.750	4.553,90
15	WIRELESS	4.083	171	11.454,00	90	LAN	1.481	27	4.387,20
16	FREQUENCY	4.551	455	11.439,70	91	RECEIVER	1.699	151	4.353,70
17	FIGURE	7.325	3.331	11.299,00	92	DELAY	2.313	690	4.340,50
18	BASED	8.448	5.193	10.804,80	93	VALUES	3.008	1.540	4.332,70
19	ROUTING	3.542	40	10.690,50	94	ARCHITECTURE	2.581	998	4.325,80
20	LAYER	4.425	569	10.604,80	95	PARAMETERS	1.841	261	4.319,50
21	MOBILE	4.341	526	10.529,90	96	FRAME	2.327	716	4.313,30
22	INPUT	4.347	709	9.868,60	97	FUNCTIONS	2.505	943	4.252,30

23	MODEL	5.895	2.290	9.860,30	98	CONNECTION	2.463	908	4.222,10
24	INTERNET	4.504	910	9.651,10	99	MODE	2.128	553	4.204,20
25	INTERFACE	3.526	207	9.557,70	100	PATH	2.700	1.210	4.196,70
26	BANDWIDTH	3.119	20	9.551,10	101	SOURCE	3.120	1.781	4.195,40
27	PACKET	3.577	251	9.485,10	102	THE	374.598	1.264.637	4.188,60
28	CIRCUIT	3.932	525	9.348,50	103	DOMAIN	1.807	280	4.153,30
29	ACCESS	5.999	2.696	9.312,50	104	OSPF	1.284	0	4.027,40
30	OUTPUT	4.139	771	9.075,40	105	LOOP	1.766	283	4.026,10
31	SERVER	3.574	362	8.963,20	106	WAVELENGTH	1.352	24	4.010,00
32	USERS	4.493	1.144	8.941,90	107	EXAMPLE	6.379	8.009	3.992,60
33	PERFORMANCE	5.686	2.712	8.539,10	108	S	6.315	7.895	3.976,30
34	DIGITAL	3.595	488	8.516,70	109	LOGIC	1.920	448	3.934,60
35	SOFTWARE	4.575	1.412	8.470,90	110	JAVA	1.344	34	3.912,60
36	SIMULATION	2.817	73	8.189,30	111	RESULTS	3.712	2.943	3.882,50
37	DEVICES	3.430	476	8.086,60	112	SECTION	3.912	3.341	3.822,90
38	CONTROL	7.124	5.343	7.818,40	113	SINGLE	4.368	4.181	3.808,40
39	VOLTAGE	2.945	220	7.743,20	114	COMPONENT	1.981	555	3.808,10
40	OPTICAL	2.822	164	7.658,10	115	NUMBER	7.759	11.407	3.799,70
41	TRAFFIC	4.345	1.615	7.420,80	116	ANALYSIS	3.153	2.115	3.781,60
42	APPLICATION	4.478	1.824	7.316,60	117	MESSAGES	1.836	433	3.749,30
43	ALGORITHM	2.799	229	7.264,20	118	SWITCH	2.075	687	3.731,10
44	INFORMATION	9.161	9.807	7.049,80	119	ANALOG	1.264	27	3.715,00
45	NODE	2.822	294	7.042,00	120	QOS	1.155	0	3.622,70
46	USED	11.874	15.508	7.023,70	121	VHDL	1.150	0	3.607,10
47	LINK	3.853	1.268	6.946,70	122	ANTENNA	1.242	34	3.597,90
48	E	6.607	5.286	6.856,90	123	DISTRIBUTED	1.824	487	3.568,90
49	SIGNALS	2.893	393	6.852,30	124	LINEAR	1.590	276	3.552,30
50	TECHNOLOGY	4.969	2.794	6.749,20	125	LEVEL	5.309	6.419	3.500,30
51	FILTER	2.627	247	6.671,30	126	SHOWN	3.507	2.930	3.498,30
52	MULTIPLE	3.046	585	6.621,90	127	MPLS	1.112	0	3.487,90
53	CURRENT	5.064	3.101	6.492,40	128	MODELS	2.255	1.024	3.479,80
54	COMMUNICATIONS	3.144	779	6.315,60	129	SOLUTIONS	1.824	523	3.475,60
55	CHANNEL	3.212	858	6.284,10	130	GSM	1.109	4	3.427,30
56	FIG	3.702	1.440	6.187,30	131	DEFINED	2.355	1.188	3.422,20
57	TRANSMISSION	2.544	308	6.171,40	132	SPECTRUM	1.499	240	3.418,10
58	PACKETS	2.308	166	6.099,60	133	ERROR	2.044	796	3.413,90
59	NODES	2.361	232	5.951,60	134	LINUX	1.128	16	3.377,10
60	PROCESS	5.949	5.015	5.886,10	135	INTERFACES	1.187	44	3.364,10
61	ROUTERS	1.891	4	5.876,00	136	POWER	5.031	6.073	3.324,30
62	WEB	2.978	791	5.838,20	137	PROGRAMMING	1.560	326	3.308,90
63	PROTOCOLS	1.996	52	5.800,40	138	CHANNELS	1.569	335	3.306,50
64	VALUE	5.063	3.633	5.761,30	139	PHASE	2.139	983	3.279,20
65	CIRCUITS	2.122	131	5.717,70	140	REMOTE	1.770	546	3.277,40
66	MANAGEMENT	4.890	3.393	5.714,70	141	ALLOWS	2.173	1.038	3.259,70
67	SERVICE	7.085	7.291	5.707,00	142	CELL	2.089	933	3.252,90
68	DEVICE	2.801	680	5.665,50	143	OBJECT	2.313	1.236	3.244,90
69	FIBER	1.869	19	5.658,80	144	CABLE	1.715	501	3.244,60
70	IMPLEMENTATION	2.734	633	5.617,10	145	AUTHENTICATION	1.082	16	3.234,20
71	ETHERNET	1.897	37	5.601,50	146	N	3.408	3.002	3.232,70
72	COMPONENTS	2.727	636	5.589,50	147	SERVERS	1.208	79	3.231,90
73	PROCESSING	2.770	686	5.565,10	148	SECURITY	3.578	3.340	3.202,20
74	CAN	25.288	53.245	5.525,10	149	VIRTUAL	1.499	310	3.189,50
75	SERVICES	6.161	5.742	5.524,20	150	INTEGRATED	1.634	444	3.176,70

Tabla 8.6 *Keywords* positivas de Teleco

Ante los resultados obtenidos, se advierten inmediatamente notables diferencias entre el listado de *keywords* y el listado de frecuencia obtenido en el primer procesamiento del corpus (Apéndice I). Al comparar las 100 primeras palabras de cada lista, se observa una escasa presencia de palabras relacionadas con las telecomunicaciones en el listado de frecuencia (*network/s*, *system/s*, *information*, *design*, *control*, *service/s*, *signal*, *user*, *access*, *process*, *performance*, *IP* y

applications); mientras que en el listado de *keywords*, todas las palabras están relacionadas con la temática excepto *is*, *each* y *can*. Por otro lado, en el listado de frecuencia predominan las palabras funcionales, al contrario que en el listado de *keywords*, donde casi todas las palabras son de contenido y están relacionadas con las telecomunicaciones en mayor o menor grado. Entre ellas, es posible distinguir cualitativamente términos técnicos restringidos a la disciplina como *IP*, *bandwidth*, *ethernet*, *TCP*, *ATM* y *LAN*. Además, entre las siguientes 50 *keywords*, aumenta la presencia de términos restringidos, los cuales incluso se pueden identificar cuantitativamente, gracias al valor de su frecuencia en el corpus general (alta frecuencia en el corpus técnico y frecuencia cero en el corpus general: *OSPF*, *QOS*, *VHDL*, y *MPLS*).

Las palabras que obtienen los valores más altos de relevancia o *keyness* resumen y especifican el contenido temático del corpus. El hecho de que se hayan identificado estadística y empíricamente los descriptores temáticos del corpus técnico, es decir, de la ingeniería de telecomunicación, resulta ser lo suficientemente razonable como para considerar que las *keywords* sean el punto de partida de un análisis más exhaustivo.

8.3.2 Distribución de las *keywords* por áreas de conocimiento.

El procedimiento seguido para identificar cómo se distribuyen las palabras clave es análogo al empleado en el apartado 7.1.2 para examinar la distribución de las palabras por las diferentes áreas de conocimiento. La técnica practicada consiste en el contraste de un listado general de *keywords* con cada una de las nueve listas de *keywords*, correspondientes a las áreas de conocimiento y las dos intensificaciones.

El listado general de palabras clave es el obtenido en la sección anterior (Apéndice IV), procedente de la comparación del corpus técnico con el corpus de lengua general que establece la norma. Sin embargo, para hallar las *keywords* de cada área, se ha utilizado el corpus técnico como lenguaje de referencia, con la intención de hacer resaltar las palabras distintivas de cada área. Se parte del supuesto de que el comportamiento estadístico de las palabras comunes entre un lenguaje específico y uno de sus componentes temáticos será similar, mientras que las palabras características del subdominio revelarán valores significativos de *keyness*. De esta manera, sería posible hacer destacar los términos específicos y las palabras generales que adquieren un significado especializado dentro del dominio.

Cabe mencionar la consulta realizada a varios expertos en las áreas de conocimiento sobre el resultado de los listados. Tras generar una lista de palabras clave tomando como referencia el corpus técnico y otra con el corpus general, se muestran a los expertos oportunos para que elijan cuál de ellas representa mejor, a su juicio, el

contenido semántico de su campo. La mayoría de ellos opta por el listado generado cuando el corpus técnico establece la norma.

El proceso seguido para hallar las *keywords* incluye los siguientes pasos, que se desarrollan para cada área de conocimiento:

1. Se extraen del corpus técnico las muestras correspondientes al área de conocimiento cuyas *keywords* se van a identificar.
2. Se halla el listado de frecuencia del área en cuestión y, por otro lado, el listado de frecuencia del resto de muestras del corpus.
3. Se obtiene el listado de palabras clave con las listas precedentes.
4. Se extraen las palabras clave cuyo margen de error es igual a cero ($p = 0$) y se introducen en una hoja de Excel. Este criterio de selección se ha adoptado como recurso para reducir el volumen de datos manejado y concentrar las palabras más significativas. El siguiente cuadro indica el número de *keywords* halladas con margen de error 0 en cada área.

Áreas de conocimiento	Nº de <i>keywords</i> $p = 0$
Electrónica	1.054
Arq. Computadoras	501
Telemática	799
Señal	769
Materiales	561
Empresa	636
Sistemas	550
Esp. Señal	936
Esp. Telemática	903

Cuadro 8.4 Volumen de *Keywords* $p = 0$.

5. Finalmente, se contrasta el listado de *keywords* individual con el listado de *keywords* de referencia cuyo margen de error también es cero (5.834 *keywords*).

En el Apéndice IV se encuentra la tabla de los resultados obtenidos sobre la distribución de las palabras clave por áreas temáticas, de donde procede la siguiente muestra seleccionada al azar:

N	KEYWORDS	Telc. Freq.	Lacell Freq.	Electr	Arq. Comp	Telmat	Señal	Materl	Empres	Sist	Esp. Señal	Esp. Telmt
1	NETWORK	16.649	1.686	-	-	3	-	-	-	-	-	802
2	DATA	14.613	2.787	-	-	-	-	-	-	-	-	802
11	LAYER	4.425	569	-	-	-	-	-	-	-	-	802
24	SIMULATION	2.817	73	1	2	-	4	-	-	7	-	-
48	COMPONENTS	2.727	636	1	2	-	4	-	-	7	-	-
56	ALGORITHMS	1.777	102	-	-	-	4	-	-	-	-	-
62	WAVELENGTH	1.352	24	1	-	-	-	5	-	-	801	-
66	SWITCH	2.075	687	1	-	3	-	-	-	-	-	802
68	QOS	1.155	0	-	-	-	-	-	-	7	-	802
72	LINEAR	1.590	276	1	-	-	4	-	-	7	-	-
82	CABLE	1.715	501	-	-	-	-	-	-	-	-	802
83	AUTHENTICATION	1.082	16	-	-	3	-	-	-	-	801	802
85	VPN	1.007	5	-	-	3	4	-	6	-	-	-
89	BROADBAND	1.049	33	-	-	-	-	-	-	-	-	802
94	MODULATION	908	19	-	-	-	4	-	-	-	801	-
95	SWITCHES	1.107	144	-	-	3	-	-	-	-	-	802
96	DESTINATION	1.311	323	-	-	-	-	-	-	-	-	802
97	NETWORKING	1.030	94	-	-	-	-	-	-	-	-	802
98	MULTICAST	837	0	-	-	-	-	-	-	-	-	802
99	VENDORS	1.004	81	-	-	3	-	-	-	-	-	802
104	PROCESSES	1.912	1.138	-	-	-	-	-	6	7	-	-
105	CISCO	840	14	-	-	3	-	-	-	-	-	802
107	RELAY	926	64	-	-	3	-	-	-	-	-	802
109	DIRECTORY	1.251	348	-	-	3	-	-	-	-	-	802
111	TOPOLOGY	811	14	-	-	-	-	-	-	-	-	802
113	CARRIER	1.102	215	1	-	-	-	-	-	-	801	-
114	SEMICONDUCTOR	845	33	1	-	-	-	5	-	-	-	-
115	LATENCY	777	7	-	-	-	4	-	-	-	-	802
116	THROUGHPUT	806	20	-	-	-	-	-	-	-	-	802
118	CHIP	1.229	353	1	2	-	-	-	-	-	-	-
120	DISCRETE	949	108	-	-	-	4	-	-	-	-	-
124	GHZ	734	2	-	-	-	4	-	-	-	801	-
126	FUZZY	851	57	1	-	-	-	-	-	7	-	-
127	SILICON	897	89	1	-	-	-	5	-	-	-	-
131	SCTP	703	0	-	-	-	-	-	-	-	-	802
138	GRAPH	1.016	224	1	2	-	4	5	-	-	-	-
143	ISDN	699	14	-	-	-	-	-	-	-	-	802
153	OSI	636	15	-	-	-	-	-	-	-	-	802
193	STATIC	808	235	-	-	-	-	-	-	7	-	-
222	OVERHEAD	792	280	-	-	-	-	-	-	7	-	-
318	OVERVIEW	508	141	-	-	-	-	-	-	-	-	-
351	KERNEL	332	26	-	-	-	-	-	-	7	-	-
372	SUBSET	328	32	-	-	-	-	-	-	7	-	-
397	LOOPS	378	84	-	-	-	-	-	-	7	-	-
399	SCALABLE	278	11	-	-	-	-	-	-	-	-	-
406	BROWSER	314	38	-	-	-	-	-	-	-	-	-
410	CONCURRENCY	244	1	-	-	-	-	-	-	7	-	-
457	HTML	323	75	-	-	-	-	-	-	-	-	-
461	TOPOLOGIES	217	2	-	-	-	-	-	-	-	-	-
465	COUPLED	450	229	-	-	-	-	-	-	-	-	-
497	COMPUTATIONS	214	9	-	-	-	-	-	-	-	-	-
499	DEPLOY	266	45	-	-	-	-	-	-	-	-	-
512	SIMULTANEOUSLY	496	342	-	-	-	-	-	-	-	-	-

523	THEOREM	278	64	-	-	-	-	-	-	-	-	-
668	QUANTUM	234	64	1	2	-	-	5	-	7	-	-
780	CVD	121	0	-	-	-	-	5	-	-	-	-
943	ANALYZER	99	0	-	-	-	-	5	-	-	-	-
1163	SCHOTTKY	83	2	1	-	-	-	-	-	-	-	-
1173	WAFERS	93	8	1	-	-	-	-	-	-	-	-
5581	EMBEDDEDNESS	9	0	-	-	-	-	-	6	-	-	-
5582	PAETZOLD	9	0	-	-	-	-	-	6	-	-	-
5597	CBDTPA	9	0	-	-	-	-	-	6	-	-	-
5656	JAD	9	0	-	-	-	-	-	6	-	-	-
5737	MANAGER'S	47	62	-	-	-	-	-	6	-	-	-
5771	GEEK	29	26	-	-	-	-	-	-	-	801	-
5772	OVERSPENDING	19	10	-	-	-	-	-	-	-	801	-
5787	RECIRCULATION	13	3	-	-	-	-	-	-	-	801	-
5794	NOAA	14	4	-	-	-	-	-	-	-	801	-
5795	SYBASE	14	4	-	-	-	4	-	-	-	-	-
5806	INTUITIONS	18	9	-	-	3	-	-	-	-	-	-
5811	AUCTION	119	250	-	-	-	-	-	-	-	801	-
5815	TAMPER	22	15	-	-	-	-	-	-	-	801	-
5821	SPUR	69	117	-	-	-	-	-	-	-	801	-
5822	IMPULSIVE	31	31	-	-	-	4	-	-	-	-	-

Tabla 8.7 Distribución de *keywords* por áreas de conocimiento.

La primera observación destacable es el comportamiento de *network* y *data*. Las dos palabras que obtienen mayor puntuación en el test de verosimilitud, es decir, las dos palabras más significativas del corpus de telecomunicaciones, sólo son *keywords* en dos y en un área de conocimiento respectivamente. De hecho, ninguna palabra es clave en todas las secciones. El valor máximo de distribución lo alcanzan sólo cuatro palabras (*simulation*, *components*, *graph*, *quantum*), presentes en cuatro áreas. La mayoría de las *keywords* del listado de referencia (3.509) sólo adquieren tal categoría en una única sección. Incluso, se advierte un elevado número de *keywords* (1.767) en el listado de referencia que no aparecen entre las *keywords* más relevantes ($p = 0$) de los listados individuales. Por lo tanto, se infiere que tales palabras se convierten en *keywords* como resultado de la suma de todas sus incidencias en el conjunto del corpus. Respecto a los valores de distribución restantes, 487 palabras son clave en dos áreas y 67 en tres. El siguiente cuadro ilustra la proporción de *keywords* en función de los valores de distribución obtenidos:

Nº de <i>Keywords</i>	Valor de distribución	Porcentaje
3.509	1	60,1%
487	2	8,3%
67	3	1,1%
4	4	0,06%
1.767	0	30,2%

Cuadro 8.5 Proporción de *keywords*.

Áreas de conocimiento	Keywords únicas
Electrónica	570
Arq. Computadoras	223
Telemática	558
Señal	357
Material	150
Empresa	200
Sistemas	238
Esp. Señal	577
Esp. Telemática	636

Cuadro 8.6 *Keywords* únicas de cada área.

8.3.3 Key-keywords.

Como se ha definido en apartados anteriores, una palabra es clave cuando su frecuencia es inusualmente alta o baja en comparación con una determinada norma. En el caso de las palabras clave positivas, cuya frecuencia es significativamente alta, el programa WordSmith permite profundizar un poco más en su comportamiento, facilitando el acceso a las *key-keywords*. Éstas se definen como palabras que son a su vez *keywords* en varios textos que componen un corpus: “*Key-keywords are those words that are key in a large number of texts of a given type*” (Scott, 1997:237).

Para hallar las *key-keywords*, el programa compara la frecuencia de las palabras en los dos corpus, teniendo en cuenta además, la frecuencia de las palabras clave dentro del corpus de estudio. La información disponible en la Tabla 8.8 indica el número de archivos en los que una palabra ha resultado ser clave, y este mismo valor expresado como porcentaje del número total de archivos. El índice de relevancia es proporcional al número de textos en los que se repite una palabra clave. Por ejemplo, la palabra *system* es clave en 533 archivos de los 1.654 que componen el corpus. En otras palabras, en el 32,22% del corpus, la palabra *system* se muestra como clave. En el Apéndice IV se encuentra el listado completo de las *key-keywords*.

N	Key-keyword	of 1,654	AS %	N	Key-keyword	of 1,654	AS %
1	DATA	672	40.63	101	PACKET	184	11.12
2	NETWORK	594	35.91	102	BETWEEN	184	11.12
3	SYSTEMS	583	35.25	103	MODELS	184	11.12
4	USING	541	32.71	104	LINK	182	11.00
5	SYSTEM	533	32.22	105	TOOLS	182	11.00
6	IS	509	30.77	106	INTEGRATED	181	10.94
7	APPLICATIONS	504	30.47	107	PATH	180	10.88
8	BASED	476	28.78	108	ERROR	179	10.82
9	THE	441	26.66	109	THUS	179	10.82

10	NETWORKS	429	25.94	110	WEB	178	10.76
11	TECHNOLOGY	396	23.94	111	TECHNIQUES	177	10.70
12	USED	371	22.43	112	CONNECTION	175	10.58
13	FIGURE	368	22.25	113	RESEARCH	175	10.58
14	DESIGN	365	22.07	114	MODE	174	10.52
15	SOFTWARE	359	21.70	115	TYPES	173	10.46
16	CAN	356	21.52	116	PROGRAM	172	10.40
17	DEVICES	351	21.22	117	PRODUCTS	172	10.40
18	INFORMATION	348	21.04	118	BE	172	10.40
19	SIGNAL	348	21.04	119	IMPLEMENTED	171	10.34
20	USER	345	20.86	120	N	171	10.34
21	ACCESS	343	20.74	121	DOMAIN	171	10.34
22	PERFORMANCE	335	20.25	122	SOLUTIONS	169	10.22
23	MODEL	331	20.01	123	FIELD	168	10.16
24	DIGITAL	322	19.47	124	PROTOCOLS	168	10.16
25	USERS	314	18.98	125	COMPUTER	168	10.16
26	FREQUENCY	314	18.98	126	CONFIGURATION	167	10.10
27	APPLICATION	313	18.92	127	PAPER	167	10.10
28	INTERNET	312	18.86	128	SPEED	166	10.04
29	MULTIPLE	311	18.80	129	AN	166	10.04
30	E	306	18.50	130	X	164	9.92
31	CONTROL	301	18.20	131	OPERATING	163	9.85
32	INTERFACE	294	17.78	132	SET	162	9.79
33	EACH	290	17.53	133	VOLTAGE	161	9.73
34	USE	290	17.53	134	ADDRESS	160	9.67
35	SERVICES	289	17.47	135	ENGINEERING	158	9.55
36	INPUT	288	17.41	136	LINEAR	158	9.55
37	SERVICE	282	17.05	137	T	157	9.49
38	IMPLEMENTATION	281	16.99	138	RATE	156	9.43
39	PROTOCOL	279	16.87	139	ROUTING	156	9.43
40	OUTPUT	278	16.81	140	STRUCTURE	155	9.37
41	PROVIDES	274	16.57	141	SUPPORT	154	9.31
42	PROCESS	263	15.90	142	PROCESSES	154	9.31
43	DEVICE	263	15.90	143	THESE	153	9.25
44	COMMUNICATIONS	262	15.84	144	G	153	9.25
45	ARE	261	15.78	145	C	152	9.19
46	IP	256	15.48	146	DISTRIBUTED	152	9.19
47	HIGH	255	15.42	147	DEVELOPMENT	152	9.19
48	CIRCUIT	255	15.42	148	AND	152	9.19
49	THIS	255	15.42	149	SPECIFIC	151	9.13
50	COMPONENTS	249	15.05	150	ORDER	151	9.13
51	CURRENT	246	14.87	151	REQUIREMENTS	151	9.13
52	SHOWN	246	14.87	152	DELAY	151	9.13
53	S	245	14.81	153	NOISE	151	9.13
54	RESULTS	244	14.75	154	PRODUCT	150	9.07
55	ANALYSIS	242	14.63	155	COST	150	9.07
56	LAYER	241	14.57	156	PHASE	150	9.07
57	EXAMPLE	240	14.51	157	DIFFERENT	150	9.07
58	SINGLE	239	14.45	158	NODE	149	9.01
59	TECHNOLOGIES	238	14.39	159	PROPOSED	148	8.95
60	MANAGEMENT	238	14.39	160	SHOWS	147	8.89
61	TRANSMISSION	235	14.21	161	FEATURES	146	8.83
62	COMMUNICATION	235	14.21	162	DUE	146	8.83
63	LEVEL	232	14.03	163	ALGORITHMS	146	8.83

64	OF	231	13.97	164	CAPABILITIES	146	8.83
65	VALUE	230	13.91	165	REQUIRED	145	8.77
66	PROCESSING	229	13.85	166	COMPONENT	145	8.77
67	BANDWIDTH	229	13.85	167	METHODS	144	8.71
68	FOR	226	13.66	168	DYNAMIC	144	8.71
69	FUNCTION	226	13.66	169	A	144	8.71
70	WIRELESS	224	13.54	170	BIT	143	8.65
71	SOURCE	221	13.36	171	EQUIPMENT	143	8.65
72	SIGNALS	220	13.30	172	CORRESPONDING	142	8.59
73	STANDARD	219	13.24	173	CHANNELS	142	8.59
74	POWER	219	13.24	174	CONNECTED	141	8.52
75	FUNCTIONS	214	12.94	175	OPERATION	140	8.46
76	HARDWARE	214	12.94	176	B	140	8.46
77	NUMBER	211	12.76	177	CHARACTERISTICS	140	8.46
78	PARAMETERS	210	12.70	178	LOGIC	140	8.46
79	PROVIDE	209	12.64	179	VIA	140	8.46
80	TRAFFIC	206	12.45	180	PACKETS	139	8.40
81	VALUES	206	12.45	181	FOLLOWING	139	8.40
82	USES	202	12.21	182	GENERATION	139	8.40
83	SECTION	202	12.21	183	BEHAVIOR	139	8.40
84	FIG	201	12.15	184	PROGRAMMING	139	8.40
85	SIMULATION	199	12.03	185	WE	138	8.34
86	ALLOWS	198	11.97	186	CUSTOMERS	138	8.34
87	MOBILE	197	11.91	187	MESSAGES	138	8.34
88	SERVER	197	11.91	188	R	138	8.34
89	TYPE	197	11.91	189	RANGE	138	8.34
90	DEFINED	196	11.85	190	NODES	137	8.28
91	METHOD	195	11.79	191	TYPICALLY	137	8.28
92	CODE	195	11.79	192	SECURITY	137	8.28
93	LOW	193	11.67	193	OBTAINED	135	8.16
94	CHANNEL	191	11.55	194	ELECTRONIC	135	8.16
95	ARCHITECTURE	190	11.49	195	BAND	135	8.16
96	SUCH	189	11.43	196	ROUTER	134	8.10
97	ALGORITHM	188	11.37	197	VIRTUAL	134	8.10
98	CIRCUITS	188	11.37	198	SPECIFICATION	133	8.04
99	SOLUTION	187	11.31	199	INFRASTRUCTURE	133	8.04
100	APPROACH	185	11.19	200	TOOL	133	8.04

Tabla 8.8 *Key-keywords*.

CAPÍTULO 9

FAMILIAS DE PALABRAS TÉCNICAS

9.1 Introducción.

El tamaño de vocabulario necesario para poder entender determinados tipos de textos ha sido objeto de diversas investigaciones. Entre ellas, Hirsh y Nation (1992, en Nation y Waring, 1997) centran su estudio en las novelas escritas para lectores adolescentes, cuyos autores tienden a utilizar un lenguaje simplificado, debido a que los destinatarios no son adultos. Los resultados de su estudio demuestran que, aunque existen más de 54.000 familias de palabras en inglés, y aunque los hablantes nativos cultos del inglés conocen alrededor de 20.000, se necesita un número mucho menor, entre 3.000 y 5.000 familias de palabras, como base para la comprensión lectora. Incluso, un tamaño más reducido, entre 2.000 y 3.000 familias de palabras, es suficiente para utilizar el lenguaje de manera productiva. El mismo número de palabras cubrirá una fracción diferente dependiendo del tipo de texto, siendo los textos académicos aquellos que requieren un vocabulario más denso. El Cuadro 9.1 muestra la porción de texto que cubren los diferentes tamaños de vocabulario en las novelas para adolescentes (Hirsh y Nation, 1992).

La unidad utilizada para medir el volumen del vocabulario requerido en cada ocasión es la familia de palabra, entendida en el sentido que le confiere Nation: “*A word family consists of a base word and all its derived and inflected forms*” (Bauer and Nation, 1993:253).

Tamaño del vocabulario	Cobertura de texto
2.000 palabras	90%
2.000 + nombres propios	93,7%
2.600 palabras	96%
5.000 palabras	98%

Cuadro 9.1 Tamaño de vocabulario y porcentaje de cobertura en novelas para adolescentes.

Conocer el porcentaje de cobertura de palabras en un texto resulta de gran utilidad. Una tasa del 80% significa que una de cada 5 palabras es desconocida, es decir, alrededor de dos palabras por línea; con el 90%, no se conoce una palabra de cada 10

(alrededor de una palabra por línea), y con el 95% se encuentra una palabra desconocida cada 20 (alrededor de una palabra cada dos líneas). De acuerdo con Laufer (1989, en Nation, 2001) un porcentaje del 95% es suficiente para permitir una comprensión razonable del texto. Sin embargo, Hirsh y Nation (1992) sugieren que, para alcanzar una lectura ágil y placentera, es conveniente poseer un vocabulario que cubra el 98-99% del texto. De este modo sólo se desconocería una palabra de cada 50 o 100.

El número de familias de palabras requeridas para cubrir un determinado porcentaje de palabras en un texto depende de varios factores: el tipo de texto (novelas, periódicos, textos académicos, conversaciones informales, etc.), la longitud del texto y la homogeneidad del texto (si trata de temas diferentes, si está escrito por uno o varios autores, etc.).

Sutarsyah, Nation y Kennedy (1994, en Nation y Waring, 1997) comparan un solo texto de economía de gran longitud con un corpus de longitud similar, pero formado por muestras de textos cortos académicos y variados. El texto de economía contiene 5.438 familias de palabras, mientras que el académico alcanza las 12.744 familias de palabras. De este modo queda comprobado que, cuando el contenido del texto se centra en un único tema, el tamaño del vocabulario requerido para la comprensión es menor.

Los resultados obtenidos respecto a textos académicos indican que sería necesario poseer un vocabulario de alrededor de 4.000 familias de palabras para alcanzar una cobertura del 95%. Tal vocabulario debería estar compuesto por “2.000 *high-frequency general service words, about 570 general academic words and 1.000 or more technical words, proper nouns and low-frequency words*” (Nation, 2001:147).

Si se trasladan estos hallazgos al ámbito del inglés profesional y académico de la ingeniería de telecomunicación, se obtendría lo siguiente: las 2.000 familias de palabras más frecuentes, junto con el vocabulario académico, conseguirían cubrir alrededor del 90% de las palabras del texto. Si además se añaden las 1.000 familias de palabras correspondientes al vocabulario técnico y las palabras de baja frecuencia, el lector se acercaría al umbral del 95% de cobertura requerido para alcanzar una comprensión razonable del texto. De este modo, si se toma la *General Service List* de West como representativa de las 2.000 familias de palabras más frecuentes del inglés y a ella se añade la lista de vocabulario académico de Coxhead, faltaría identificar cuáles son las 1.000 familias de palabras técnicas más frecuentes de las telecomunicaciones y las palabras de baja frecuencia, según se refleja en el siguiente cuadro:

Tipo de vocabulario	Familias de palabras	Cobertura	Lista
Vocabulario de alta frecuencia	2.000	± 80%	General Service List, West, 1953
Vocabulario académico	570	± 9%	Academic Word List, Coxhead, 2000
Vocabulario técnico	1.000	± 5%	¿?
Vocabulario de baja frecuencia	Indefinido	± 5%	----

Cuadro 9.2 Cobertura del vocabulario.

9.2 Proceso de selección de las familias de palabras técnicas.

9.2.1 Criterios preliminares.

A la luz de los resultados obtenidos en los análisis precedentes, se pretenden establecer los criterios que permitan seleccionar las mil familias de palabras técnicas imprescindibles en el vocabulario de la ingeniería de telecomunicaciones. Estos se identificarán en función de los siguientes criterios preliminares:

1. Las palabras que pertenezcan a las mil familias de palabras técnicas deben ser palabras de contenido. Las palabras funcionales quedan excluidas tras aplicar la lista de corte correspondiente.
2. Las palabras que formen parte de las mil familias de palabras técnicas no deben aparecer ni en el listado de West (1953), ni en el de Coxhead (2000). Sin embargo, se ha decidido no prescindir de las palabras académicas y generales más frecuentes, en caso de que alguna de ellas adquiriera un significado especializado o muestre ser significativamente relevante en telecomunicaciones.
3. Las palabras que integren las mil familias de palabras técnicas deben ser *keywords* positivas. No es suficiente que una palabra sea muy frecuente en un dominio específico, sino que debe ser estadísticamente destacable en comparación con su uso en el lenguaje general.
4. Las palabras que se incluyan en las mil familias de palabras técnicas deben de ser valoradas como términos específicos de acuerdo con los criterios propuestos por Chung (2003). De lo contrario, se estudiará el caso en búsqueda de las posibles razones que expliquen dicho comportamiento y así comprobar el carácter específico.

El listado de palabras siguiente constituye la base de donde extraer las mil familias de palabras técnicas. En el Apéndice V se encuentra la lista de palabras que revela los datos necesarios para comprobar qué palabras cumplen con los requisitos planteados, y que permite, a la vez, reunir a los miembros de cada familia. La siguiente tabla ofrece una muestra del listado base y del tipo de información que contiene:

N	KEYWORD	Telec. Frec.	FRT	Lacell Frec.	FRL	Ratio	Término, Chung	Keyness Log L.	P valor
1	NETWORK	16.649	0,30086533	1.686	0,00802234	37,5034384	NO	41.784,60	0
2	DATA	14.613	0,26407262	2.787	0,01326113	19,9132855	NO	31.852,20	0
3	SIGNAL	7.022	0,12689509	641	0,00305001	41,6047867	NO	17.922,60	0
4	SYSTEMS	9.479	0,17129572	3.000	0,01427463	12,0000145	NO	17.377,70	0
5	IP	5.239	0,09467436	20	9,5164E-05	994,852983	ESPECIFICO	16.182,10	0
6	NETWORKS	5.832	0,10539051	463	0,00220305	47,8384421	NO	15.204,90	0
7	SYSTEM	12.624	0,22812926	8.707	0,04142972	5,50641507	NO	14.831,60	0
8	USER	6.292	0,11370321	903	0,00429666	26,4631457	NO	14.725,80	0
9	PROTOCOL	4.742	0,08569304	139	0,00066139	129,564867	ESPECIFICO	13.677,70	0
10	DESIGN	7.701	0,13916535	3.313	0,01576395	8,8280786	NO	12.237,80	0
11	APPLICATIONS	5.414	0,0978368	934	0,00444417	22,0146548	NO	12.117,70	0
12	ROUTER	3.910	0,0706579	25	0,00011896	593,987427	ESPECIFICO	11.974,40	0
13	USING	9.214	0,16650689	5.576	0,02653177	6,27575451	NO	11.914,70	0
14	WIRELESS	4.083	0,0737842	171	0,00081365	90,6825603	ESPECIFICO	11.454,00	0
15	FREQUENCY	4.551	0,08224146	455	0,00216498	37,9870827	NO	11.439,70	0
16	FIGURE	7.325	0,13237063	3.331	0,01584959	8,35167336	NO	11.299,00	0
17	BASED	8.448	0,15266444	5.193	0,02470938	6,17840092	NO	10.804,80	0
18	ROUTING	3.542	0,06400775	40	0,00019033	336,301705	ESPECIFICO	10.690,50	0
19	LAYER	4.425	0,07996451	569	0,00270742	29,5353086	NO	10.604,80	0
20	MOBILE	4.341	0,07844654	526	0,00250282	31,3432874	NO	10.529,90	0
21	INPUT	4.347	0,07855496	709	0,00337357	23,2854111	NO	9.868,60	0
22	MODEL	5.895	0,10652899	2.290	0,0108963	9,77662215	NO	9.860,30	0
23	INTERNET	4.504	0,08139212	910	0,00432997	18,7973875	NO	9.651,10	0
24	INTERFACE	3.526	0,06371861	207	0,00098495	64,6922813	ESPECIFICO	9.557,70	0
25	BANDWIDTH	3.119	0,05636368	20	9,5164E-05	592,278384	ESPECIFICO	9.551,10	0
26	PACKET	3.577	0,06464024	251	0,00119431	54,1234812	ESPECIFICO	9.485,10	0
27	CIRCUIT	3.932	0,07105547	525	0,00249806	28,4442646	NO	9.348,50	0
28	ACCESS	5.999	0,10840838	2.696	0,01282813	8,45083219	NO	9.312,50	0
29	OUTPUT	4.139	0,07479618	771	0,00366858	20,3883252	NO	9.075,40	0
30	SERVER	3.574	0,06458602	362	0,00172247	37,496133	NO	8.963,20	0
31	USERS	4.493	0,08119334	1.144	0,00544339	14,9159493	NO	8.941,90	0
32	PERFORMANCE	5.686	0,10275213	2.712	0,01290426	7,96265087	NO	8.539,10	0
33	DIGITAL	3.595	0,06496552	488	0,00232201	27,9781875	NO	8.516,70	0
34	SOFTWARE	4.575	0,08267517	1.412	0,00671859	12,3054332	NO	8.470,90	0
35	SIMULATION	2.817	0,05090622	73	0,00034735	146,5563	ESPECIFICO	8.189,30	0
36	DEVICES	3.430	0,06198379	476	0,00226491	27,3670302	NO	8.086,60	0
37	CONTROL	7.124	0,12873834	5.343	0,02542311	5,06383144	NO	7.818,40	0
38	VOLTAGE	2.945	0,05321932	220	0,00104681	50,8397167	ESPECIFICO	7.743,20	0
39	OPTICAL	2.822	0,05099657	164	0,00078035	65,3512148	ESPECIFICO	7.658,10	0
40	TRAFFIC	4.345	0,07851882	1.615	0,007684507	10,2178085	NO	7.420,80	0
101	OSPF	1.284	0,02320326	0	0	Inf/esp	Inf/esp	4.027,40	0

Tabla 9.1 Listado base para familias técnicas.

Como se puede apreciar en la tabla anterior, para cada una de las *keywords*, se especifica la frecuencia absoluta (Telec. Frec. y Lacell Frec.) y la frecuencia relativa (FRT y FRL) en cada uno de los corpus. Estas cifras permiten calcular el valor del ratio ($< \text{ó} > 50$), con el fin de discernir si la palabra clave se considera término específico según los criterios establecidos por Chung. La columna encabezada por “Término, Chung” presenta tres claves posibles:

- Específico: La palabra se considera término porque el ratio es > 50 .
- No: La palabra no se considera término porque el ratio es < 50 .
- Inf/esp: La palabra se considera término porque el ratio es infinito, es decir, la palabra no aparece en el corpus general.

A continuación, en la columna *Keyness Log L.*, se hace constar el índice de relevancia obtenido para cada palabra. Finalmente, se muestra el margen de error del cálculo, reflejado en la última columna P valor.

9.2.2 Familia de palabras: Agrupación de miembros.

Una vez que se dispone de las palabras clave positivas de contenido, ordenadas de forma descendente en función de su valor de relevancia, junto a la estimación del carácter específico de cada una de ellas, se procede a agrupar los diferentes miembros de cada familia.

El concepto de familia de palabras operativo en este estudio responde a la definición de Bauer y Nation, a la que se hizo una breve referencia anteriormente. No obstante, es conveniente desarrollar y explicitar este concepto, tal y como queda formulado en Bauer y Nation (1993:253):

A word family consists of a base word and all its derived and inflected forms that can be understood by a learner without having to learn each form separately. [...] The important principle behind the idea of a word family is that once the base word or even a derived word is known, the recognition of other members of the family requires little or no extra effort. Clearly, the meaning of the base in the derived word must be closely related to the meaning of the base when it stands alone or occurs in other derived forms.

Siguiendo tales directrices y según el juicio de la analista, se han agrupado los diferentes miembros de las familias, tomando como representante o cabeza de familia al miembro que obtiene mayor puntuación en el test de razón de verosimilitudes, es decir, la *keyword* más significativa representa al resto. La primera familia deriva de *network* y, una vez que se han reunido todos los miembros, se repite la operación con la segunda *keyword*, después con la siguiente que le sucede en relevancia, y así sucesivamente

hasta alcanzar mil familias de palabras. A continuación se muestran, a modo de ejemplo, las cinco primeras familias identificadas (Tabla 9.2):

Familias de palabras	F Telec	F. Gene	Ratio	Término	Keyness	P valor
1 NETWORK	16.649	1.686	37,5034384	NO	41.784,60	0
NETWORKS	5.832	463	47,8384421	NO	15.204,90	0
NETWORKING	1.030	94	41,6149977	NO	2.628,20	0
INTERNETWORK	195	1	740,585347	ESPECIFICO	599,5	0
SUBNETWORK	133	0	inf/esp	inf/esp	417,1	0
NETWORKED	141	27	19,8333398	NO	306,7	0
INTERNETWORKING	82	3	103,808544	ESPECIFICO	232,6	0
INTERNETWORKS	53	0	inf/esp	inf/esp	166,2	0
SUBNETWORKS	45	0	inf/esp	inf/esp	141,1	0
MULTINETWORK	12	0	inf/esp	inf/esp	37,6	0
FOUNDRYNETWORKS	7	0	inf/esp	inf/esp	22	0,000003
ADNETWORK	6	0	inf/esp	inf/esp	18,8	0,000014
NETWORKMESSENGER	4	0	inf/esp	inf/esp	12,5	0,000397
2 DATA	14.613	2.787	19,9132855	NO	31.852,20	0
DATAGRAM	181	0	inf/esp	inf/esp	567,7	0
DATAGRAMS	116	0	inf/esp	inf/esp	363,8	0
METADATA	53	1	201,2873	ESPECIFICO	156,7	0
DATATYPE	20	0	inf/esp	inf/esp	62,7	0
DATASHEET	19	0	inf/esp	inf/esp	59,6	0
DATATYPES	15	0	inf/esp	inf/esp	47	0
DATASHEETS	12	1	45,5744829	NO	31,1	0
DATAPRO	8	0	inf/esp	inf/esp	25,1	0,000001
DATAIN	7	0	inf/esp	inf/esp	22	0,000003
DATAOUT	5	0	inf/esp	inf/esp	15,7	0,000075
DATASHEMA	5	0	inf/esp	inf/esp	15,7	0,000075
DATUM	19	20	3,6079799	NO	14,9	0,000113
INPUTDATA	4	0	inf/esp	inf/esp	12,5	0,000397
OUTPUTDATA	4	0	inf/esp	inf/esp	12,5	0,000397
DATARATES	4	0	inf/esp	inf/esp	12,5	0,000397
DATATAC	4	0	inf/esp	inf/esp	12,5	0,000397
3 SIGNAL	7.022	641	41,6047867	NO	17.922,60	0
SIGNALS	2.893	393	27,9573747	NO	6.852,30	0
SIGNALING	601	22	103,751001	ESPECIFICO	1.705,00	0
SIGNALLING	170	88	7,33680123	NO	243,2	0
SIGNALED	21	17	4,69149089	NO	21,6	0,000003
SIGNALTNOISE	5	0	inf/esp	inf/esp	15,7	0,000075
4 SYSTEM	12.624	8.707	5,50641507	NO	14.831,60	0
SYSTEMS	9.479	3.000	12,0000145	NO	17.377,70	0
SUBSYSTEM	170	5	129,127702	ESPECIFICO	490,1	0
SUBSYSTEMS	126	9	53,1702301	ESPECIFICO	333,3	0
LINKSYS	44	0	inf/esp	inf/esp	138	0
SYSTEMC	43	0	inf/esp	inf/esp	134,9	0
MICROSYSTEMS	63	17	14,0744727	NO	122,8	0
FILESYSTEM	22	0	inf/esp	inf/esp	69	0
SYSTEMATIC	142	256	2,106633	NO	46,4	0
INTERSYSTEM	8	0	inf/esp	inf/esp	25,1	0,000001
MICROSYSTEM	8	1	30,3829886	NO	19,3	0,000011
LESYSTEM	6	0	inf/esp	inf/esp	18,8	0,000014

SYSTEMVIEW	4	0	inf/esp	inf/esp	12,5	0,000397
VSYSTEM	4	0	inf/esp	inf/esp	12,5	0,000397
5 IP	5.239	20	9,5164E-05	ESPECIFICO	16.182,10	0
IPS	17	14	6,6615E-05	NO	17,2	0,000034
IPBASED	5	0	0	inf/esp	15,7	0,000075
IPCENTER	4	0	0	inf/esp	12,5	0,000397

Tabla 9.2 Familias de palabras.

9.2.3 Marcado de palabras académicas y generales más frecuentes.

En esta etapa de la investigación, se dispone de las mil familias de palabras más significativas del inglés de la ingeniería de telecomunicación, independientemente de su naturaleza. Sin embargo, aún no se han distinguido cuáles componen el vocabulario especializado de este dominio.

De acuerdo con los criterios establecidos anteriormente, las familias técnicas no deben aparecer en los listados de West (1953) y Coxhead (2000), aunque en principio no quedan descartadas en caso de que pudieran adquirir un carácter especializado o ser especialmente significativas. Consecuentemente, se procede a realizar un cotejo entre los diferentes listados con el fin de marcar qué familias pertenecen a cada categoría. De este modo, se adhiere la etiqueta de “Académica” o “West” a las formas que aparecen en la lista del vocabulario académico o general más frecuente, según corresponda. Si ninguna de las *keywords* es localizada en las listas de referencia, pero sí se encuentra el lema de donde derivan, éste se añade junto a la etiqueta del miembro más significativo de la familia (Ver ejemplo en Tabla 9.3, familia 324 *RESISTOR* / *West: resist*).

Una tercera etiqueta: “West técnica”, designa las palabras que se hallan en la lista del vocabulario científico-técnico básico de Flood y West: *Word-List for the Writing of Popular Science and Technology* (West, 1953). Esta lista aparece como complemento de la *General Service List* y fue elaborada con la finalidad de exponer a los lectores sin formación académica al lenguaje de la ciencia. De acuerdo con sus creadores, el listado comprende el vocabulario necesario para entender “*all ordinary scientific and technical subjects within the limits of non-specialized study*” (West, 1953:583). Entre las 425 palabras que componen la lista, 69 son términos altamente especializados y moderadamente necesarios, pero son incorporados porque su uso resulta indispensable. La presencia de estas palabras en la lista original se distingue gráficamente en letras mayúsculas, y del mismo modo se marcará la etiqueta correspondiente en la lista de familias de palabras: “WEST TÉCNICA”.

De nuevo por motivos de espacio, se presenta una muestra de las familias marcadas en la Tabla 9.3. El documento completo se encuentra disponible en el Apéndice V:

Familias de palabras	F. Telec	F. Gene	Ratio	Término	Keyness	P valor
1 NETWORK / Académica	16.649	1.686	37,5034384	NO	41.784,60	0
NETWORKS / Académica	5.832	463	47,8384421	NO	15.204,90	0
NETWORKING / Académica	1.030	94	41,6149977	NO	2.628,20	0
INTERNETWORK	195	1	740,585347	ESPECIFICO	599,5	0
SUBNETWORK	133	0	inf/esp	inf/esp	417,1	0
NETWORKED / Académica	141	27	19,8333398	NO	306,7	0
INTERNETWORKING	82	3	103,808544	ESPECIFICO	232,6	0
INTERNETWORKS	53	0	inf/esp	inf/esp	166,2	0
SUBNETWORKS	45	0	inf/esp	inf/esp	141,1	0
MULTINETWORK	12	0	inf/esp	inf/esp	37,6	0
FOUNDRYNETWORKS	7	0	inf/esp	inf/esp	22	0,000003
ADNETWORK	6	0	inf/esp	inf/esp	18,8	0,000014
NETWORKMESSENGER	4	0	inf/esp	inf/esp	12,5	0,000397
4 SYSTEM / West	12.624	8.707	5,50641507	NO	14.831,60	0
SYSTEMS	9.479	3.000	12,0000145	NO	17.377,70	0
SUBSYSTEM	170	5	129,127702	ESPECIFICO	490,1	0
SUBSYSTEMS	126	9	53,1702301	ESPECIFICO	333,3	0
LINKSYS	44	0	inf/esp	inf/esp	138	0
SYSTEMC	43	0	inf/esp	inf/esp	134,9	0
MICROSYSTEMS	63	17	14,0744727	NO	122,8	0
FILESYSTEM	22	0	inf/esp	inf/esp	69	0
SYSTEMATIC / West	142	256	2,106633	NO	46,4	0
INTERSYSTEM	8	0	inf/esp	inf/esp	25,1	0,000001
MICROSYSTEM	8	1	30,3829886	NO	19,3	0,000011
LESYSTEM	6	0	inf/esp	inf/esp	18,8	0,000014
SYSTEMVIEW	4	0	inf/esp	inf/esp	12,5	0,000397
VSYSTEM	4	0	inf/esp	inf/esp	12,5	0,000397
5 IP	5.239	20	994,852983	ESPECIFICO	16.182,10	0
IPBASED	5	0	inf/esp	inf/esp	15,7	0,000075
IPCENTER	4	0	inf/esp	inf/esp	12,5	0,000397
IPS	17	14	4,61170363	NO	17,2	0,000034
6 USER / West	6.292	903	26,4631457	NO	14.725,80	0
USING	9.214	5.576	6,27575451	NO	11.914,70	0
USERS	4.493	1.144	14,9159493	NO	8.941,90	0
USED / West	11.874	15.508	2,90791532	NO	7.023,70	0
USE / West	10.255	14.866	2,61988386	NO	5.144,70	0
USES	2.302	1.296	6,74591433	NO	3.123,50	0
USAGE	727	288	9,58699337	NO	1.204,00	0
MULTIUSER	366	2	695,010865	ESPECIFICO	1.124,00	0
REUSE	327	25	49,6761864	NO	856,9	0
USEFUL / West	1.263	2.066	2,32173975	NO	507,7	0
USERNAME	58	1	220,276667	ESPECIFICO	172,2	0
USABILITY	101	45	8,52411625	NO	157,5	0
USABLE	96	44	8,28626962	NO	147,4	0
REUSED	59	21	10,6702162	NO	102,8	0
REUSABLE	47	10	17,8500058	NO	99,1	0
REUSABILITY	29	1	110,138334	ESPECIFICO	82,7	0
REUSING	27	2	51,2712933	ESPECIFICO	71,1	0
UNUSED	87	105	3,14680953	NO	57,5	0
USENET	34	16	8,07048135	NO	51,4	0
USUALLY / West	1.427	4.354	1,24473257	NO	49,6	0

UNUSABLE	25	9	10,5496488	NO	43,3	0
SINGLEUSER	9	0	inf/esp	inf/esp	28,2	0
USEFULNESS / West	66	106	2,36471374	NO	27,5	0
USERNAMES	7	0	inf/esp	inf/esp	22	0,000003
USENIX	6	0	inf/esp	inf/esp	18,8	0,000014
NEWUSERS	5	0	inf/esp	inf/esp	15,7	0,000075
USEABLE	11	6	6,96276822	NO	15,2	0,000095
USERINPUT	4	0	inf/esp	inf/esp	12,5	0,000397
USERLIST	4	0	inf/esp	inf/esp	12,5	0,000397
ENDUSER	4	0	inf/esp	inf/esp	12,5	0,000397
OVERUSED	9	6	5,69681036	NO	10,8	0,000992
7 PROTOCOL	4.742	139	129,564867	ESPECIFICO	13.677,70	0
PROTOCOLS	1.996	52	145,779917	ESPECIFICO	5.800,40	0
MULTIPROTOCOL	73	0	inf/esp	inf/esp	229	0
8 DESIGN / Académica	7.701	3.313	8,8280786	NO	12.237,80	0
DESIGNED / Académica	1.822	2.425	2,85349512	NO	1.046,60	0
DESIGNS / Académica	706	746	3,59423424	NO	551,3	0
DESIGNERS / Académica	461	329	5,32164048	NO	526,7	0
DESIGNING / Académica	357	207	6,54995588	NO	475	0
DESIGNER / Académica	327	621	1,99984647	NO	94,4	0
CODESIGN	23	0	inf/esp	inf/esp	72,1	0
REDESIGN	39	39	3,79787358	NO	32,4	0
10 ROUTER	3.910	25	0,00011896	ESPECIFICO	11.974,40	0
ROUTING / Académica	3.542	40	0,00019033	ESPECIFICO	10.690,50	0
ROUTERS	1.891	4	1,90E-05	ESPECIFICO	5.876,00	0
ROUTE / Académica	1.437	1.015	0,00482958	NO	1.655,50	0
ROUTES / Académica	891	419	0,00199369	NO	1.348,40	0
ROUTED / Académica	345	39	0,00018557	NO	848	0
ROUTABLE	32	0	0	inf/esp	100,4	0
REROUTING	22	0	0	inf/esp	69	0
ROUTABILITY	18	0	0	inf/esp	56,5	0
NONROUTABLE	14	0	0	inf/esp	43,9	0
WASTEROUTE	11	0	0	inf/esp	34,5	0
ALLSPFROUTERS	11	0	0	inf/esp	34,5	0
REROUTE	13	3	1,43E-05	NO	26,7	0
ALLDROUTERS	8	0	0	inf/esp	25,1	0,000001
XROUTER	6	0	0	inf/esp	18,8	0,000014
UNROUTABLE	6	0	0	inf/esp	18,8	0,000014
REROUTED	9	3	1,43E-05	NO	16,1	0,000059
ENROUTE	5	0	0	inf/esp	15,7	0,000075
TRACEROUTE	5	0	0	inf/esp	15,7	0,000075
LAMBDAROUTER	4	0	0	inf/esp	12,5	0,000397
40 FILTER / West técnica	2.627	247	40,3927688	NO	6.671,30	0
FILTERS	1.207	108	42,4447538	NO	3.089,60	0
FILTERING	575	37	59,0210083	ESPECIFICO	1.541,40	0
FILTERED	99	75	5,01319312	NO	107,7	0
PREFILTER	6	0	inf/esp	inf/esp	18,8	0,000014
MICROFILTER	5	0	inf/esp	inf/esp	15,7	0,000075
PREFILTERING	5	0	inf/esp	inf/esp	15,7	0,000075
309 SHARED / West: share	1.221	985	4,70782095	NO	1.257,20	0
SHARING	627	670	3,55412945	NO	483,1	0
PRESHARED	11	0	inf/esp	inf/esp	34,5	0
SHAREGET	4	0	inf/esp	inf/esp	12,5	0,000397
PROSHARE	4	0	inf/esp	inf/esp	12,5	0,000397

TIMESHARING	4	0	inf/esp	inf/esp	12,5	0,000397
324 RESISTOR / West: resist	427	18	90,094001	ESPECIFICO	1.196,90	0
RESISTORS	324	10	123,051104	ESPECIFICO	931	0
RESISTANCE / WEST TÉCNICA	686	687	3,79234538	NO	569,4	0
RESISTANCES	48	7	26,0425617	NO	111,9	0
RESISTIVE	59	6	37,3457568	NO	147,8	0
RESISTIVITY	51	8	24,2114441	NO	116,9	0

Tabla 9.3 Marcado de las familias.

Tanto en la tabla anterior como en el Apéndice V, se puede apreciar que las formas marcadas como académicas o generales, suelen concordar con la valoración de Chung. De un total de 1.100 formas etiquetadas, 1.070 se reconocen como palabras no específicas, siendo el porcentaje de coincidencia de un 97%.

No obstante, se observa el caso contrario, es decir, existen formas catalogadas como académicas o generales, consideradas a la vez específicas. Se han detectado 30 etiquetas que no concuerdan con la estimación de Chung. Este comportamiento podría ser un indicio del significado especializado que una palabra ha adquirido o, simplemente, puede que se utilice más en este ámbito por la temática sobre la que versa y la modalidad escrita del lenguaje recopilado. El siguiente cuadro contiene las formas evaluadas como específicas, de las cuales 25 son marcadas como académicas y 5 como generales:

Académicas: *Routing, simulation, dynamically, computation, computed, compute, computes, schematic, prioritized, prioritization, prioritizes, aggregation, automates, summarization, appends, refines, optionally, estimation, estimations, conformance, normalized, normalised, normalizes, inferencing.*

Generales: *packet, receivers, sender, resistor, ampere.*

Cuadro 9.3 Etiquetadas general/académica, valoradas como específicas.

En tercer lugar, se advierte que no todos los miembros de una familia participan del carácter de su representante. A veces, aunque una forma pertenezca a una familia académica o general, no va acompañada de la etiqueta correspondiente. Tales resultados encuentran una doble explicación: Por un lado, se hallan las formas que no están catalogadas en las listas de West y Coxhead, pero que resultan ser específicas aplicando el método de Chung. Estas palabras suelen ser formas compuestas o derivadas del lema base del cabeza de familia y están provistas de un significado claramente especializado, bien técnico o semitécnico. Entre ellas, 349 formas pertenecen a familias académicas y 602 a generales. Algunos ejemplos son: *internetwork, subsystem, phasor, wavelet, interoperability, mappings, codesign, router, wasteroute, photocurrent, runtime, encoder, parameterized, connectivity, multimode, intradomain, multistreaming, payload, etc.*

Por otro lado, se encuentran los miembros de familias académicas o generales sin etiquetar, que no se consideran específicos. Tal comportamiento se produce con más frecuencia en palabras generales que en académicas (703 y 268 respectivamente), debido a las propias características del listado de West y del sistema de marcado empleado. En la *General Service List*, no figuran las diferentes flexiones del mismo lema y sólo aparecen algunas de sus formas en los ejemplos, mientras que en la lista del vocabulario académico, el lema despliega sus diferentes formas. Cabe recordar que la condición establecida para etiquetar una forma en las familias de palabras, es que dicha forma se debe encontrar en el listado de referencia general o académico.

No obstante, en el presente estudio se da por supuesto que las diferentes flexiones también participan de la naturaleza del lema que las representa, a no ser que obtengan la puntuación requerida para valorarlas como términos específicos. En cualquier caso, la desambiguación del carácter específico o general de una palabra exige su estudio en los diferentes contextos donde aparece y activa su significado. Una vez más, es necesario puntualizar que la clasificación automática de términos se realiza en función de la morfología de las palabras, y el reconocimiento de secuencias de caracteres que efectúa el ordenador no posibilita la distinción de formas homógrafas ni de significados.

El resultado del marcado pone de manifiesto la incidencia de las palabras académicas y de las generales más frecuentes entre las palabras clave. Con el fin de apreciar su presencia globalmente de una forma más clara, se han dividido las 1.000 familias en series de cien expresando su ocurrencia numéricamente:

Familias de palabras	Familias académicas		Familias generales más frecuentes		Resto	
1 – 100	25	2,5 %	35	3,5 %	40	4 %
101 – 200	21	2,1 %	25	2,5 %	54	5,4 %
201 – 300	16	1,6 %	19	1,9 %	65	6,5 %
301 – 400	25	2,5 %	23	2,3 %	52	5,2 %
401 – 500	26	2,6 %	22	2,2 %	52	5,2 %
501 – 600	16	1,6 %	20	2 %	64	6,4 %
601 – 700	11	1,1 %	14	1,4 %	75	7,5 %
701 – 800	18	1,9 %	22	2,1 %	60	6 %
801 – 900	14	1,4 %	14	1,4 %	72	7,2 %
901 – 1000	14	1,4 %	12	1,2 %	74	7,4 %
Total	187	18,7 %	206	20,5 %	608	60,8 %

Cuadro 9.4 Incidencia de familias.

En cada una de las series, ordenadas de mayor a menor índice de relevancia, se muestra el número de familias generales y académicas encontradas y el porcentaje que representa. Al poder observar su incursión en los distintos rangos, se advierte cómo su presencia no es constante y tiende a disminuir conforme desciende el índice de relevancia.

Entre las mil familias de palabras más significativas del corpus de telecomunicaciones, se desvela que un 39,2% pertenece a familias académicas y generales más frecuentes (18,7% y 20,5% respectivamente). En cambio, al variar la unidad de medida, de familias de palabras a formas, se obtienen resultados distintos. Las mil familias engloban un total de 5.381 miembros, de los cuales un 19,8% se han etiquetado como formas académicas o generales. Además, el porcentaje se eleva al 28,3%, si se añaden las flexiones no etiquetadas de los lemas académicos y generales, valoradas a la vez como no específicas (70 formas académicas y 388 generales).

El análisis de los datos precedentes permite concluir que cerca de un tercio de las palabras más significativas del corpus de telecomunicaciones no son de carácter especializado, sino que forman parte del vocabulario académico e incluso del vocabulario de alta incidencia del lenguaje general. Este hecho no significa que el resto de palabras sean todas específicas, aunque sí cabe esperar una mayor presencia de términos especializados. La delimitación del vocabulario académico y general más frecuente allana el camino hacia la búsqueda de los términos específicos más relevantes.

9.3 Resultado de la clasificación de familias en el corpus de telecomunicaciones.

9.3.1 Familias de palabras técnicas.

Una vez que se ha reconocido el vocabulario académico y general más frecuente, éste se separa del listado base de familias de palabras, el cual queda considerablemente reducido y fragmentado en tres. De las 1000 familias iniciales, 608 se mantienen como candidatas a términos especializados. Esto implica que se debe continuar agrupando *keywords* en familias y reproducir el procedimiento anterior de marcado y clasificación, hasta conseguir las mil familias de palabras técnicas.

Prosiguiendo con la investigación, el resto de familias candidatas a términos se someten a revisión, donde se advierten los siguientes tipos de comportamientos:

Familias calificadas como NO específicas.

En primer lugar, es notoria la presencia de familias, compuestas por uno o varios miembros, en las que todas las formas son valoradas como no-términos según los

criterios propuestos por Chung (2003). No obstante, destacan algunas palabras que cualitativamente se estiman propias de un subdominio de las telecomunicaciones y, por tanto, es conveniente someterlas a un análisis más detallado.

Entre ellas se encuentran algunos nombres propios de empresas o marcas de productos tecnológicos, cuyo reconocimiento y uso está mundialmente extendido, razón por la cual no obtienen un valor lo suficientemente alto para ser consideradas términos. Tal es el caso de *Microsoft, IBM, Toshiba, Netscape, Intel, Ericsson* y *Nortel*.

Por otro lado, se hallan palabras de origen técnico que, debido a la globalización y difusión del conocimiento, abundan en el lenguaje general: *www (world wide web), GPS (global positioning system), AC (alternating current), DC (direct current), RAM (random access memory), transceiver, pixel, laptop* y *SIM (subscriber identity module), HTTP (hypertext transfer protocol), RADAR (radio detection and ranging)*. Estas palabras, junto al resto de miembros de sus familias, admiten ser incluidas como integrantes del vocabulario especializado.

Dos particularidades caracterizan a 34 familias: el hecho de estar formadas por un solo miembro y ser siglas o abreviaturas a la vez. Éstas se analizarán junto al resto de siglas en sus correspondientes contextos, puesto que es necesario conocer su forma extendida para determinar su naturaleza, a excepción de *et, al, vs, etc* y *km*.

De este modo, de las 151 familias no específicas iniciales, 111 se añaden al listado de vocabulario general clave utilizado en el inglés de la ingeniería de telecomunicación, 11 se incorporan al listado de vocabulario específico y 29 se reagrupan con las siglas.

Palabras etiquetadas como “WEST TÉCNICA”.

Las palabras que adquieren esta etiqueta pertenecen a la lista de vocabulario científico-técnico básico de Flood y West (1953), caracterizadas por ser términos especializados de uso moderado pero imprescindible. Por esta misma razón, *frequency, circuit, volt, solution, spectrum, power, cell, conductor, electron, electrode, generate, equation, resistance, oxide, meter, ion* y *ampere*, permanecen como términos específicos. Curiosamente, sólo la forma *ampere* es valorada como término. El resto de palabras, en su condición de formas individuales, no son lo suficientemente frecuentes en este dominio especializado en comparación con el lenguaje general para lograr la calificación de término, aunque sí lo son varios miembros de sus respectivas familias. La familia de *spectrum* ilustra esta conducta:

Familias de palabras	F. Telec	F. Gene	Ratio	Término	Keyness	P valor
98 SPECTRUM / WEST TÉCNICA	1.499	240	23,7208854	NO	3.418,10	0
SPECTRAL	578	17	129,127702	ESPECIFICO	1.666,40	0
SPECTRA	168	24	26,585115	NO	393,5	0
SPECTROSCOPY	45	21	8,13830052	NO	68,4	0
SPECTRALLY	13	0	inf/esp	inf/esp	40,8	0

SPECTROMETER	21	6	13,2925575	NO	40,1	0
BISPECTRUM	8	0	inf/esp	inf/esp	25,1	0,000001
SPECTROMETERS	11	2	20,8883047	NO	24,3	0,000001
TRANSPECTRUM	7	0	inf/esp	inf/esp	22	0,000003
SPREADSPECTRUM	6	0	inf/esp	inf/esp	18,8	0,000014
LEONARDOSPECTRUM	6	0	inf/esp	inf/esp	18,8	0,000014
SPECTRORADIOMETERS	5	0	inf/esp	inf/esp	15,7	0,000075
HYPERSPECTRAL	5	0	inf/esp	inf/esp	15,7	0,000075
SPECTROSCOPIC	9	4	8,54521555	NO	14	0,000178
POLYSPECTRA	4	0	inf/esp	inf/esp	12,5	0,000397
760 AMP	267	133	7,62430259	NO	390,8	0
AMPS	147	17	32,8404362	NO	359,8	0
OPAMP	43	0	inf/esp	inf/esp	134,9	0
AMPERE / WEST TÉCNICA	41	3	51,9042722	ESPECIFICO	108,1	0
IAMPS	27	0	inf/esp	inf/esp	84,7	0
AMPERES	21	1	79,7553451	ESPECIFICO	58,2	0
MICROAMPS	8	0	inf/esp	inf/esp	25,1	0,000001
OPAMPS	8	0	inf/esp	inf/esp	25,1	0,000001

Tabla 9.4 Familias etiquetadas como WEST TÉCNICA.

Formas abreviadas.

Desde el punto de vista formal, cabe destacar la concurrencia de palabras aparentemente simples, cuyo análisis muestra una formación de origen complejo mediante procesos de truncación: siglación, acronimia y abreviación. Las formas extendidas son identificables por medio de la herramienta de concordancia, capaz de mostrar cada uno de los contextos en los que aparece una determinada palabra.

La mayoría de las formas complejas detectadas, representantes de familia, comparten una valoración de término específico, sucediendo lo contrario en tan sólo 15 formas. Se trata de formas de origen especializado cuyo uso se ha extendido (*SMS*, *PDF*, *e-mail*, *HTLM*, *ISO*, *XP*) o bien son representantes de una familia en la que el resto de miembros son específicos (*IC*, *DD*, *KB*, *LMS*, *ASIC*, *CNC*, *MIPS*, *SSL*, *MOSFET*), e incluso se pueden cumplir las dos propiedades a la vez (*SMS*, *PDF*, *XP*). En cualquier caso, las formas extendidas corroboran la calidad de término de las formas abreviadas.

Los procesos de truncación dan lugar a formas complejas denominadas abreviaturas, acrónimos y siglas. Las abreviaturas son formas, normalmente fijadas por consenso, que reproducen el segmento inicial de una palabra: *Sec* (*second*), *amp* (*ampere*), *khz* (*kilohertz*), *mhz* (*megahertz*), *ID* (*identification*), *mux* (*multiplexer*), etc... Los acrónimos son palabras formadas por la combinación de segmentos de un sintagma desarrollado, por ejemplo *VHDL* es el acrónimo que representa la combinación de *VHSIC* (*very high speed integrated circuit*) y *HDL* (*hardware description language*), *Comint* (*communication intelligence*), *GaAs* (*gallium arsenide*), etc. Finalmente, las siglas son unidades formadas por la combinación de las iniciales de varias palabras que

constituyen una expresión más larga. El número de siglas halladas (Cuadro 9.5) supera con creces al de abreviaturas y acrónimos. A continuación se aporta una muestra de las siglas junto a su correspondiente forma extendida.

IP (internet protocol)	ATM (asynchronous transfer mode)
TCP (transmission control protocol)	LAN (local area network)
OSPF (open shortest path first)	MPLS (multi-protocol label switching)
QOS (quality of service)	GSM (global system for mobile communications)
VPN (virtual private network)	GPRS (general packet radio service/system)
LSA (link-state advertisement)	DSP (digital signal processing)
FPGA (field programmable gate array)	HP (high-pass)
WLAN (wireless local area network)	MAC (media access control)
WAN (wide area network)	SI (silicon)
CMOS (complementary metal oxide semiconductor)	SHWP (shared wavelet path)
ITU (international telecommunications union)	PID (proportional integral derivative) algorithm
GIS (geographical information systems)	XDSL (various digital subscriber line)
WEP (wired equivalent privacy)	MIMO (multiple-input multiple output)
IEEE (institute of electrical and electronic engineering)	MOSFET (metal-oxide semiconductor field-effect transistor)
Otras siglas:	
<p><i>CDMA, SCTP, RF, DB, ISDN, OSI, PC, CPU, WDM, MBPS, NM, VOIP, LDAP, SNMP, BGP, DCE, FTP, DVB, XML, SPS, DNS, ACK, UDP, VLSI, IC, KBPS, RSVP, PMI, NS, LSR, PN, NSSA, PPP, VER, IOS, HTML, ADSL, FCC, NMS, TTL, UMTS, FEC, TMN, SNA, VLAN, DMCS, IGRP, PBX, RMI, DHCP, LSPS, UML, ISP, ICT, IPX, PDU, CPE, MIB, ARP, DD, API, JVM, MPEG, FDDI, PDF, SDSL, ICMP, STD, MN, VTP, ARR, CSP, DFT, SDL, GUI, IIR, ITIL, PPRP, WAP, URI, PSTN, MPI, HRM, TDM, CRC, KB, SNR, XP, RTT, QPSK, LMS, ASIC, ISO, GNU, CDR, LISP, CVD, HDL, CWND, NTP, TLS, EDA, SCWP, SDH, LEDS, ISM, NBMA, SN, DIGE, RTP, NOC, CNC, MIPS, DWT, DAC, VRF, DTE, OFDM, SMC, SVM, PSAP, VCC, SSL, FIFO, CN, AMS, SISO, ASP, HLR, CMT, GBPS, NGSS, RTS, DWDM, IFIP, QAM, YN, DIP, PE, SLAS, DCT, NFS, URL, VDD, CSMA, LRE, PDAS, RCS, IDS, RMS, PVCS, SMDS, NSA, APPN, TC, UDRP, AAL, ISI, SCL, CW, VLR, TM, RIP, NS, FIR, NT, VR, BS, ADA, OC, ISM, FDR, PCB, VC, PE, RS, ANSI, PCI, GB, EV, MV, MTU, CE.</i></p>	

Cuadro 9.5 Siglas

Familias en las que todos los miembros son valorados como específicos.

Un total de 158 familias de palabras están compuestas por miembros lo suficientemente frecuentes en el corpus específico en relación con el general, para ser catalogados como términos. Entre ellas se pueden reconocer fácilmente nombres propios de empresas estrechamente relacionadas con el sector (*Cisco, Xilinx, Agilent, Alcatel, Lucent, Roke: Roke Manor Research*); nombres de programas, arquitecturas, sistemas y lenguajes informáticos (*Java, Labview, Matlab, Linux, Promela: Process or*

Protocol Meta Language, Corba: Common Object Request Broker Architecture, Sonet: Synchronous Optical Network, Docsis: Data Over Cable Service Interface Specification, DECnet: Digital Equipment Corporation network); y apellidos que califican un tipo de operación en honor a la persona que la formuló: *Nyquist (theorem), Eifel (algorithm), Fourier (series, transform), Rayleigh (scattering coefficient).*

La siglación también otorga un carácter distintivo al siguiente grupo de familias de miembros específicos: *WEP, RMI, LSPS, UML, ISP, IPX, PDU, CPE, MIB, APP, MSC, JVM, MPEG, CSP, DFT, SKL, IIR, ITIL, WAP, URI, HRM, QPSK, MIMO, CVD, TLS, APPN, SCWP, SDM, LEDS, RTP, DWT, VRF, FTE, OFDM, SMC, DLSW, SVM, MUX, PSAP, VCC FIFO, SISO ASP, XDSL, CMT, GBPS, NGSS, RTS, DWDM, IFIP, QAM, YN, MTU, SLAS, VCD, DCT, NFS, VDD, CSMA, LRE, PDAS, RCS, IDS, VLR, HDL, EDA, TC, UDRP, SCL, AAL, CW, CWND, NTP, CDR, LISP, LAN, VHDL, MPLS, LSAS MHZ, FPGA, WLAN, ISDN, CMOS, LDAP, BGP, DCE, DVB, SPS, IGRP, PSTN, TDM, GNU, SNR, RTT, TCP, WDM, ACK, UDP, VLSI, PMI, PN, NSSA, PPP, ADSL, FCC, NMS, TTL, TMN, SNA, DMCS, VLAN, PBX, MPI.*

Finalmente, el resto de familias quedan catalogadas, desde una perspectiva formal, como términos claramente especializados, pues todos los miembros satisfacen cada uno de los requisitos previamente establecidos: *protocol, multicast, latency, throughput, cache, impedance, bluetooth, diode, microstrip, doppler, lookup, octects, crosstalk, ingress, firewall, timeslot, recursive, applet, handover, cosine, dialog, unicast, netlist, diffraction, vertex.*

Familias en las que más de la mitad de los miembros son valorados como específicos.

El hecho de que el número de miembros específicos supere al número de miembros estimados no específicos, permite aceptar a la familia entera como parte del vocabulario técnico: *IP, wireless, bandwidth, software, traffic, algorithm, node, web, Ethernet, configuration, architecture, analog, antenna, spectrum, interfaces, power, cable, authentication, satellite, modulation, DSP, director, topology, carrier, global, chip, fuzzy, vector, RF, DB, client, interference, hash, buffer, bits, optimization, matrix, synchronization, equation, static, size, cluster, Unix, embedded, step, encryption, proxy, scalability, Gaussian, multiplexing, magnetic, hop, threshold, Boolean, roaming, KHz, oxide, cryptographic, orthogonal, provisioning, jitter, telework, photonic, duplex, exponential, dipole, debugging, doped, spice, solaris, amp, ISO, convolution, smart, scattering, granularity, aliasing, meters, internationalization, shell, pipelined, Google, log.*

Resto de familias.

En este último grupo se incluyen aquellas familias que no manifiestan ninguno de los comportamientos anteriores y se distinguen por las siguientes características:

- a) Familias en las que la mitad de los miembros son valorados específicos: *mobile, optical, transmission, vendors, propagation, laser, case, attenuation, pulse, backup, peer, generic, correlation, benchmark, literals, metal, sinusoidal, end.*
- b) Familias en las que menos de la mitad de los miembros son valorados como específicos: *digital, information, multiple, function, distributed, relay, synthesis, execution, amplifier, electrical, default, token, backbone, subscriber, array, verification, maximum, bytes, click, segment, external, fault, collaborative, active, overview, equivalent, capacitor, dispersion, convergence, grid, query, thermal, designated, prototype, diffusion, fabrication, billing, inverting, minimum, dial, decoder, snack, oscillator, polarization, mask, entry, hackers, extensions, velocity, disk, emitter, dissipation, zeros, junction, sec, tunneling, quantum, mathematical, distortion, hub, interception, resonant, suite, checksum, navigation, profiling, replication.*
- c) Familias en las que sólo un miembro es valorado como específico: *cost, multimedia, storage, part, forwarding, feedback, broadcast, response, deployment, iteration symmetric, bridging, radiation, empirical, encapsulation, collision, exchange, scanning, installation, degradation, silicon, modem, gain, paper, coefficients, voip, metrics, STD, robust, desktop, knowledge, stub, gam, outage, coaxial, copper, heuristic.*

La siguiente gráfica muestra, a modo de resumen, las proporciones de los diferentes comportamientos hallados entre las familias de palabras candidatas a términos técnicos, siendo la familia de palabras la unidad de medida adoptada.

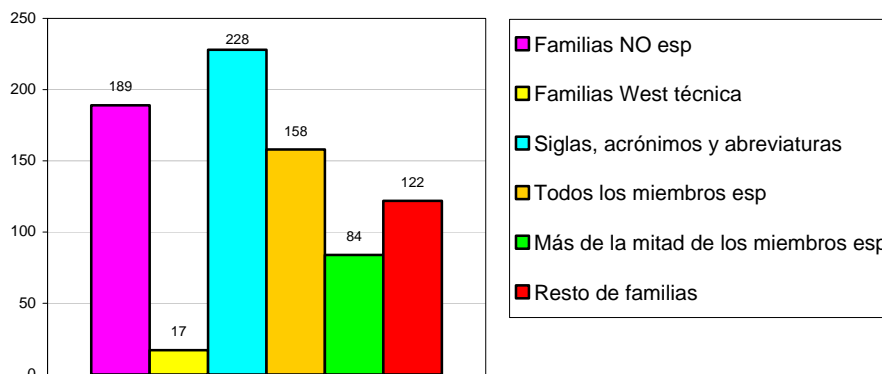


Figura 9.1 Características de las familias candidatas a términos.

9.3.2 Familias de palabras generales.

Una segunda fracción del listado base de familias de palabras corresponde al conjunto de familias generales más frecuentes, al cual se agregan las 111 familias procedentes de las 608 iniciales candidatas a términos especializados (Apéndice V). Estas 111 familias ya se desestimaron como técnicas, porque todos los miembros constituyentes son valorados no específicos, formal y estadísticamente, así que resta señalar los comportamientos observados en las pertenecientes al listado de West:

1. **Familias en las que todos los miembros son generales:** *value, provides, results, single, rate, product, engineering, proposed, support, due, description, content, introduction, existing, properties, diagram, window, length, allow, continuous, advanced, includes, desired, chain, calculated, scenario, experimental, temperature, delivery, reduce, signature, discussed, pricing, edge, advantage, companies, presented, increases, width, impulse, given, privacy, tree, thickness, observed, industry, large, curve, coverage, sent, offers, needed, supply, actual, entire, demand, battery, deposition, excitation, wrapper, originating, correctness, pattern, improved, purpose.*
2. **Familias en las que sólo un miembro es específico:** *applications, figure, performance, current, management, address, subsection, behavior, order, neighbor, associated, probability, developed, represented, density, sample, additional, related, critical, possible, problem, password, limited, tracking, matching, common, pair, market, total, laboratory, reservation, root, manufacturing, instrumentation, smaller, shielding, driven, absorption, changes, angles, ring, repeaters, unreachable.*
3. **Familias en las que más de la mitad de los miembros son generales:** *signal, user, based, control, receiver, delay, frame, connection, loop, linear, noise, Standard, load, field, set, operating, graph, different, low, key, request, customer, sender, reference, determine, state, temporal, clock, gate, clock, note, radio, testing, sensor, directly, adoption, class, organizational, compared, expressions, tuning, equal, combinational, real, atoms, handle, commerce, actions, loss, permittivity, conditional.*
4. **Familias en las que la mitad de los miembros son generales:** *messages, programming, object, memory, speed, header, mapping, fixed, command, machine, calls, faster, failures, imaging, agent, bus.*
5. **Familias en las que más de la mitad de los miembros son específicos:** *system, internet, model, packet, server, filter, fiber, type, path, wavelength, example, number, switch, tools, parallel, host, flow, phones, table, groupware, space,*

shared, payment, simple, beam, pointer, threads, stream, distance, center, dependent, stations, plane, banking.

6. **Familias de la lista de vocabulario científico-técnico básico de Flood y West (1953):** *absorb, atom, battery, fibre, crystal, dense, deposit, develop, diagram, filter, field, graph, image, impulse, laboratory, loop, parallel, plane, sensitive, section, switch.*

La gráfica 2 ilustra los diferentes comportamientos hallados entre las familias de palabras generales más frecuentes, tomando la familia de palabras como unidad de medida:

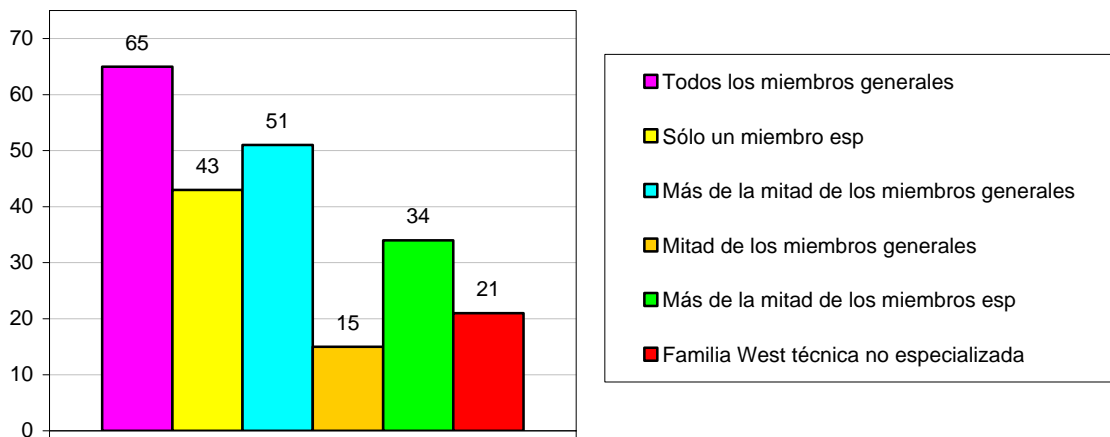


Figura 9.2 Comportamiento de las familias generales.

Los comportamientos observados permiten discernir mediante criterios formales y estadísticos, qué familias pueden haber adquirido un significado específico y cuáles mantienen su carácter general. La mayoría de los grupos clasificados responden a una conducta no especializada, dado que el número de miembros generales excede el número de miembros específicos, razón por la cual la familia en conjunto se considera general. Respecto a las palabras del listado de Flood y West, también quedan clasificadas como vocabulario general de uso altamente frecuente en un ámbito científico-técnico, aunque se hallan casos en los que los miembros clasificados como específicos superan a los generales.

Normalmente, las palabras de alta frecuencia suelen presentar un repertorio de sentidos más variado que las palabras menos frecuentes: *“the more frequent the occurrence of a word, the greater is the variety of meanings in which it is used”* (West, xi, 1953). Además, cuando los especialistas de un dominio necesitan denominar conceptos del ámbito especializado, recurren a los mecanismos disponibles en el sistema de la lengua para crear unidades léxicas nuevas. Entre ellos, los recursos semánticos están basados en la modificación del significado de una unidad, procedente

del léxico general o de otras áreas especializadas, a las que confieren un nuevo significado específico. En consecuencia, la acumulación de ocurrencias de una palabra frecuente podrá ser repetición de la misma forma, pero puede no ser simplemente reiteración del mismo sentido, sino la muestra de una complejidad mayor. Este hecho provoca que una misma forma obtenga valores similares de uso en el registro específico y en el general y que, por tanto, no manifieste diferencias significativas, cuando realmente se están computando significados distintos asociados a un significante. Por este motivo es necesario insistir en que el contexto dictamina el carácter de la unidad léxica y desambigua sentidos.

A lo largo del análisis realizado, se han detectado algunas palabras que, aun siendo propias del lenguaje general y tras haber obtenido la puntuación correspondiente, presentan ciertas dudas sobre la posibilidad de haber sufrido un cambio o adaptación del significado general a un sentido específico: *banking, bus, host, linear, log, mapping, memory, noise, program, radio, escenario, shell, signal, tuning, window*. Estas palabras se someterán a un análisis detallado, en el que se estudiarán los contextos oportunos, con el fin de comprobar si en el dominio de las telecomunicaciones se han utilizado con un sentido diferente al registrado en la *General Service List*, y si es así, identificar sus patrones léxicos.

9.3.3 Familias de palabras académicas.

Entre las mil familias de palabras más significativas del inglés de la ingeniería de telecomunicación, 190 pertenecen al vocabulario académico (Apéndice V). Las diferentes formas registradas en el listado académico de Coxhead que también se encuentran entre las *keywords* base, suelen concordar con la estimación de palabra no específica, salvo las 25 excepciones presentadas en el Cuadro 9.3. Las 349 formas restantes, es decir, aquellas que no aparecen en la lista de vocabulario académico pero son miembros relacionados, adquieren la calificación de término.

A diferencia de las palabras generales, las académicas no requirieron un acuciente análisis contextual más detallado para probar si se activan significados diferentes, puesto que una de las características esenciales del vocabulario académico es su no vinculación a un dominio en particular, aun siendo frecuentes en una amplia variedad de disciplinas. No obstante, Wang y Nation (2004) examinan la *Academic Word List* de Coxhead, con el objetivo de comprobar si alguna de sus formas está provista de varios significados lo suficientemente distinguibles entre sí como para formar familias independientes. Asimismo se comprueba, en términos de Wang y Nation (2004:291), “*if the existence of unrelated meanings for the same word form (homographs) has resulted in the inclusion of words in the list which would not be there if their clearly different meanings were distinguished*”.

El estudio demuestra que aproximadamente el 10% de las familias académicas contenían palabras homógrafas que no deberían figurar en una única familia. Tan sólo 21 de estos homógrafos ocurrían con la frecuencia suficiente para afectar al estatus de su familia, y por tanto, requerían una atención individual: *consist, issue, credit, normal, correspond, volume, attribute, project, decline, generation, objective, abstract, brief, intelligent, appreciate, offset, tense, induce, accommodate, converse* y *panel*. Finalmente, tras examinar cada caso, de las 570 familias iniciales sólo varían 9. Los cambios se reducen a 6 familias que demandan una entrada diferente y 3 quedan eliminadas (*intelligent, offset* y *panel*).

En definitiva, la lista de vocabulario académico de Coxhead es lo suficientemente fiable para asumir que las *keywords* halladas en el corpus no estarán dotadas de significados adicionales, siempre que coincidan con las formas registradas en el listado académico. No obstante, las familias de *network, router, data* y *layer*, se someterán a un análisis detallado, puesto que parecen establecer una estrecha relación con el área, además de estar formadas, en mayor parte, por miembros no etiquetados como académicos y valorados como específicos. Por lo tanto, se comprobará si adquieren un significado especializado en el dominio.

9.4 Conclusión y criterios finales.

Los resultados obtenidos tras la aplicación de los criterios preliminares propuestos para la identificación de las familias de palabras técnicas, han mostrado ser doblemente satisfactorios. No sólo se ha conseguido registrar un buen número de familias técnicas, sino que, además, el desarrollo de la aplicación ha permitido colegir los ajustes necesarios para distinguir el vocabulario especializado.

La decisión de no prescindir de las palabras académicas y generales más frecuentes ha sido acertada, puesto que se han detectado posibles usos especializados entre ellas. Después, la combinación de los criterios de los que se sirven los métodos estadísticos basados en la comparación de corpus ha sido determinante.

En primer lugar, una palabra debe ser clave, es decir, estadísticamente significativa en el lenguaje especializado en comparación con su incidencia en el lenguaje general. A continuación, se reúnen los miembros de las familias partiendo de la forma más representativa y se aplica la valoración de Chung (2003) a cada uno de los componentes.

El método propuesto por Chung ha resultado ser suficientemente fiable y válido para identificar formas especializadas. Sin embargo, las unidades que componen las listas de vocabulario son familias de palabras. En consecuencia, se ha decidido considerar familia técnica a aquella en la que todos o la mayoría de los miembros se

valoran como términos especializados. En el caso contrario donde los miembros específicos no son mayoría, se ha resuelto estimar las formas individualmente, es decir, si una forma resulta ser específica según Chung, se toma como término independientemente de la calidad del resto de los miembros de su familia. Por lo tanto, el vocabulario especializado se compondría de familias técnicas completas y de formas técnicas que no comparten el carácter especializado con el resto de su familia. Tales formas especializadas se hallarían en una situación similar a las flexiones de un lema que no participan plenamente del significado de la voz que las representa: “*Meaning is not constant across the inflected forms of a lemma*” (Stubbs, 1996:40).

De este modo, la aplicación de los criterios y los ajustes realizados demuestran ser válidos para alcanzar el objetivo planteado de identificar el listado de palabras especializadas más significativas e imprescindibles en el lenguaje de la ingeniería de telecomunicación. La lista de vocabulario especializado que se ha obtenido (Apéndice V), consta de 402 familias técnicas más 1.017 formas técnicas individuales, incluidas en las 1.000 familias de palabras más significativas del corpus de telecomunicaciones.

En cuanto a la proporción de texto que la lista de vocabulario técnico consigue cubrir, se ha calculado que el total de formas alcanza el 9,16% del corpus completo. Entre ellas se encuentran formas con distintos grados de relevancia y probabilidad de aparición, formas de alta y baja frecuencia, formas valoradas como específicas, etc... Paulatinamente, logran alcanzar el porcentaje de palabras que es necesario conocer para poder comprender un texto profesional y/o académico del dominio especializado. Cabe recordar que, de acuerdo con las conclusiones extraídas de diversos estudios realizados al respecto (Laufer, 1989; Hirsh y Nation, 1992; Sutarsyah, Nation y Kennedy, 1994; Nation y Waring, 1997; Nation, 2001), las 2.000 familias generales más frecuentes, junto con el vocabulario académico, cubren alrededor del 90% de las palabras de un texto. Del 10% restante, un 5% compete al vocabulario especializado y el otro 5% corresponde a palabras de baja frecuencia y de gran diversidad.

Sin embargo, los miembros de las 402 familias técnicas valorados como específicos no son suficientes para alcanzar el porcentaje requerido puesto que sólo llegan a un 2,89%, aunque al añadir el resto de los miembros, el porcentaje asciende a un satisfactorio 4,17%. No obstante, entre ellos se encuentran términos técnicos con una probabilidad de aparición menor que otras formas más significativas pero de menor grado de especialización. Por esta razón, es pertinente considerar las formas identificadas como específicas independientemente de la naturaleza del conjunto de su familia, de modo que el total de formas valoradas como específicas en la lista íntegra de vocabulario técnico cubre un 4,08%, del cual el 4% corresponde a aquellas formas cuyo P valor es igual a cero (Cuadro 9.6).

	Suma de frecuencias	Cobertura del corpus
Todas las formas	507.139	9,16 %
Miembros valorados específicos en familias técnicas	160.231	2,89%
Todos los miembros de las familias técnicas	441.394	4,17%
Total formas valoradas específicas	225.976	4,08%
Total formas valoradas específicas p valor = 0	221.415	4%

Cuadro 9.6 Proporción de formas y su cobertura del corpus.

Finalmente, el listado definitivo de palabras pertenecientes al vocabulario especializado de las telecomunicaciones se expondrá más adelante, tras resolver las dudas planteadas sobre el posible uso especializado de algunas formas.

PALABRAS CON PALABRAS: ANÁLISIS DETALLADO

10.1 Introducción.

Esta etapa de la investigación está enfocada hacia el análisis detallado de una muestra de palabras seleccionadas. Para cada una de ellas, se recapitula la información obtenida previamente sobre su comportamiento estadístico en relación con los factores de frecuencia, distribución y relevancia, junto a las conclusiones obtenidas al respecto. Asimismo, se explora el contexto inmediato de cada palabra con el fin de describir las relaciones sintagmáticas establecidas y, de este modo, dar cuenta del comportamiento léxico de las muestras elegidas.

De acuerdo con la corriente lingüística en la que se ubica el presente trabajo, se considera que los rasgos estadísticos del conjunto de muestras definen el sublenguaje especializado de las telecomunicaciones, en función de la variación de los elementos léxicos y gramaticales respecto al lenguaje general. De este modo, el significado de tales elementos elegidos en el discurso se interpreta tanto por lo que expresan como por lo que excluyen. Esta concepción encaja con el primer principio de interpretación de significado en los textos, propuesto por Sinclair, *the open-choice principle*: “*This is a way of seeing language text as the result of a very large number of complex choices. At each point where a unit is completed (a word, phrase, or clause), a large range of choice opens up and the only restraint is grammaticalness*” (Sinclair, 1991:109).

Por otro lado, las relaciones léxicas en el nivel sintagmático conciernen a las relaciones semánticas que se establecen entre un lexema y el resto que le acompaña, es decir, entre palabras que co-ocurren o aparecen juntas en un determinado contexto próximo. Tales relaciones están íntimamente ligadas al concepto de colocación y al segundo principio de interpretación del significado textual de Sinclair, *the idiom principle*: “*a language user has available to him or her a large number of semi-preconstructed phrases that constitute single choices, even though they might appear to be analysable into segments*” (Sinclair, 1991).

Las herramientas disponibles en WordSmith son idóneas para llevar a cabo el análisis pertinente, puesto que permiten identificar colocados y clusters partiendo de las líneas de concordancias. De acuerdo con las definiciones del propio programa, se

entiende por colocados “*the words which occur in the neighbourhood of your search word*” (Scott, 1998). Tales colocados ayudan a mostrar el significado y uso de la palabra bajo análisis. Por otro lado, los clusters o agrupaciones léxicas se definen como “*words which are found repeatedly in each others’ company [which] represent a tighter relationship than collocates*” (Scott, 1998).

El estudio de las relaciones sintagmáticas conduce a la identificación de combinaciones de palabras prefabricadas utilizadas por los especialistas en situaciones de comunicación especializada. Por lo tanto, estas combinaciones constituyen un factor de caracterización del sublenguaje especializado, además de ser fundamentales para la adecuada expresión y comprensión del conocimiento especializado.

A continuación, se ilustra con la palabra clave *bandwidth* el tipo de análisis realizado, argumentando detalladamente el desarrollo del proceso llevado a cabo. Para el resto de palabras analizadas, se aporta una ficha con la información obtenida sobre el comportamiento o patrón léxico identificado.

10.2 Desarrollo del análisis detallado: *Bandwidth*.

El análisis desarrollado consta de cuatro secciones: frecuencia, distribución, colocados y clusters. Como se ha mencionado anteriormente, las dos primeras están relacionadas con el principio de *open-choice* o libre elección, en el marco de un registro especializado; mientras que las dos últimas se corresponden con el *idiom principle*, donde se activan las restricciones que no se pueden captar mediante el ejercicio de la libre elección. Una vez que se conoce el léxico elegido en el registro y su distribución por los diferentes subdominios, los colocados y clusters muestran cómo se utiliza en el discurso de la comunidad de hablantes.

10.2.1 Frecuencia.

El factor frecuencia pone de manifiesto la elección de un elemento lingüístico en el sublenguaje de las telecomunicaciones frente a la lengua general, y muestra si tal elección es lo suficientemente incidente como para considerar término específico al elemento lingüístico. El mismo tipo de información queda reflejada para el resto de los miembros de la familia, y se añade la calificación de familia técnica, general o académica, según corresponda.

Tomando *bandwidth* como ejemplo, su comportamiento estadístico indica que es la vigésima palabra más significativa del corpus, con una puntuación de 9.551 en relevancia o *keyness*. Además, *bandwidth* resulta ser un término específico del dominio de acuerdo con los criterios propuestos por Chung (2003). Toda la familia de *band*, representada por *bandwidth*, se considera técnica, puesto que más de la mitad de los

miembros que la componen son específicos (22 miembros específicos frente a 2 generales).

KEYWORD	Frec. Telec	Frec. Lacell	Ratio	Término, Chung	KEYNESS	P valor
20. BANDWIDTH	3.119	20	592,278384	ESPECIFICO	9.551,10	0
Miembros relacionados:	F Telec	F Gene	Ratio	Término	Keyness	P valor
BROADBAND	1.049	33	120,726345	ESPECIFICO	3.010,30	0
BAND	1.760	1.630	4,10077147	NO	1.587,70	0
WIDEBAND	177	1	672,223623	ESPECIFICO	543,3	0
PASSBAND	151	0	inf/esp	inf/esp	473,6	0
BASEBAND	148	0	inf/esp	inf/esp	464,2	0
BANDGAP	146	0	inf/esp	inf/esp	457,9	0
NARROWBAND	138	0	inf/esp	inf/esp	432,8	0
STOPBAND	99	0	inf/esp	inf/esp	310,5	0
BANDWIDTHS	97	1	368,393737	ESPECIFICO	293,5	0
INFINIBAND	68	0	inf/esp	inf/esp	213,3	0
BANDS	350	469	2,83423401	NO	199	0
BANDPASS	50	3	63,2978929	ESPECIFICO	135,2	0
SIDEBAND	21	0	inf/esp	inf/esp	65,9	0
SUBBAND	20	0	inf/esp	inf/esp	62,7	0
HALFBAND	19	0	inf/esp	inf/esp	59,6	0
SUBBANDS	17	0	inf/esp	inf/esp	53,3	0
SIDEBANDS	11	0	inf/esp	inf/esp	34,5	0
PASSBANDS	10	0	inf/esp	inf/esp	31,4	0
STARBAND	7	0	inf/esp	inf/esp	22	0,000003
STOPBANDS	7	0	inf/esp	inf/esp	22	0,000003
BANDLIMITED	4	0	inf/esp	inf/esp	12,5	0,000397
INBAND	4	0	inf/esp	inf/esp	12,5	0,000397
MULTIBAND	4	0	inf/esp	inf/esp	12,5	0,000397
Familia técnica						

Tabla 10.1 Familia de *band*.

10.2.2 Distribución.

La información obtenida respecto a la distribución de una unidad léxica muestra la disposición y recurrencia de tal unidad en las diferentes secciones que integran el corpus específico de telecomunicaciones. En la ficha correspondiente (ver ejemplo más abajo), las casillas sobre la distribución indican en qué áreas aparece la palabra en cuestión y en cuáles es *keyword*. Los valores posibles para la distribución por áreas oscilan del 1 al 9, señalando entre paréntesis y con un signo negativo el número del área correspondiente donde no aparece la palabra (-5). Después, en la casilla “Distribución *keyness*”, se indica en qué área la unidad es clave. A continuación, se especifica si la palabra clave adquiere la categoría de *key-keyword*, mostrando el número de textos en los que se revela como *keyword* y la proporción que aquéllos ocupan en el corpus.

Finalmente, se adjunta una representación gráfica de la distribución de las palabras, mediante un gráfico de dispersión (*dispersion plot*) que facilita el programa WordSmith. El gráfico muestra, para cada área, el número total de formas (*words*), la frecuencia de

la palabra en el área (*hits*), el número de ocurrencias por cada mil formas (*per 1,000*) y el dibujo correspondiente a tales datos.

Siguiendo con el ejemplo adoptado, *bandwidth* aparece en las nueve áreas de conocimiento, pero sólo es palabra clave en las dos especialidades: Sistemas y redes de telecomunicación (081) y Planificación y gestión de telecomunicaciones (082). Esto significa que, aunque *bandwidth* está presente y se relaciona en mayor o menor medida con todos los subdominios de las telecomunicaciones, su incidencia sólo es significativa en dos áreas. Además, *bandwidth* es *key-keyword* en 229 archivos de los 1.654 que componen el corpus, es decir, su presencia es significativa en el 13,85% del corpus.

Distribución por áreas		Distribución Keynes		Keykeyword	Nº de textos	Porcentaje
9		081+ 082		BANDWIDTH	229	13.85%
<i>Dispersion plot</i>						
Area	Words	Hits	Per 1,000			
081 Esp.Sign.	867.175	997	1.15			
082 Esp.Tele	997.683	959	0.96			
4 Signal proc.	580.890	259	0.45			
1 Electronics	722.778	309	0.43			
3 Telematics	1.204.955	381	0.32			
7 Systems	307.662	72	0.23			
6 Business	373.043	49	0.13			
2 Ar. Comp	329.605	36	0.11			
5 Materials	101.232	8	0.08			

Tabla 10.2 Distribución de *bandwidth*.

10.2.3 Colocación y colocados significativos.

El concepto de colocación léxica engloba todas las combinaciones frecuentes de unidades léxicas. Éstas se combinan estableciendo relaciones con distinto grado de atracción, que oscilan entre la existencia de una cierta afinidad entre palabras, hasta el punto de convertirse en locuciones (*idioms*), siempre que el significado del conjunto no se desprende de la suma del significado de sus componentes.

En términos de Sinclair (1991:115): “*words appear to be chosen in pairs or groups and these are not necessarily adjacent*”. Si la presencia de dos palabras en un contexto próximo es lo suficientemente recurrente como para advertir que su co-ocurrencia no se debe a la casualidad, se constituye una colocación significativa. Dependiendo de la frecuencia y del puesto que ocupa cada palabra (nodo o colocado), se establecen dos tipos de colocación que contribuyen de forma diferente en la descripción de patrones léxicos. Por un lado, cuando los colocados presentan una frecuencia más alta que la del nodo (palabra bajo análisis), se produce la denominada *upward collocation* o colocación ascendente. En cambio, el caso contrario da lugar a la colocación descendente o

downward collocation, donde el nodo es más frecuente que el colocado. Entre los dos tipos de colocación existe una diferencia sistemática que consiste en lo siguiente: “*Upward collocation is the weaker pattern in statistical terms, and the words tend to be elements of grammatical frames, or superordinates. Downward collocation by contrast gives us a semantic analysis of a word*” (Sinclair, 1991:116).

Si la colocación descendente facilita el análisis semántico de una palabra, la identificación de los colocados significativos de una *keyword* contribuirá a desvelar el sentido con el que se ha utilizado, y a despejar las dudas que puedan surgir sobre la categoría a la que pertenece. Por consiguiente, compete hallar los colocados de la palabra bajo estudio, empleando la herramienta destinada a ello en WordSmith. El programa requiere primero visualizar en pantalla las líneas de concordancia del nodo, a partir de las cuales presenta los colocados. Retomando *bandwidth*, la *keyword* que ilustra todos los pasos seguidos en el análisis, estos son los resultados obtenidos:

N	Collocates	TOTAL	LEFT	RIGHT	L5	L4	L3	L2	L1	*	R1	R2	R3	R4	R5
1	THE	2309	1409	900	177	175	195	430	432	0	4	319	215	174	188
2	OF	1149	594	555	78	72	121	143	180	0	260	86	50	77	82
3	AND	874	380	494	70	63	75	82	90	0	153	116	59	102	64
4	TO	850	465	385	97	117	154	89	8	0	84	83	64	75	79
5	A	699	384	315	63	73	67	105	76	0	9	107	63	72	64
6	IS	604	182	422	61	44	36	39	2	0	149	66	71	65	71
7	IN	441	189	252	37	34	47	43	28	0	41	70	35	54	52
8	FOR	428	223	205	31	52	52	59	29	0	52	58	37	31	27
9	WITH	248	147	101	17	23	41	46	20	0	24	22	17	18	20
10	THAT	242	137	105	26	45	30	24	12	0	32	16	21	16	20
11	ON	208	86	122	24	14	13	28	7	0	51	18	16	15	22
12	BE	204	72	132	24	26	17	4	1	0	0	57	25	16	34
13	BY	186	61	125	10	17	12	21	1	0	21	30	27	23	24
14	THIS	182	78	104	29	14	13	9	13	0	0	30	27	22	25
15	AS	180	80	100	15	15	15	21	14	0	18	20	19	26	17
16	HIGH	168	138	30	5	4	4	6	119	0	0	12	8	6	4
17	AVAILABLE	167	104	63	2	1	2	13	86	0	37	5	12	7	2
18	NETWORK	162	107	55	14	21	15	10	47	0	1	4	9	21	20
19	CAN	154	61	93	19	23	12	7	0	0	26	12	17	18	20
20	MORE	151	111	40	15	14	17	7	58	0	1	10	8	12	9
21	ARE	148	72	76	30	22	11	7	2	0	5	21	19	16	15
22	AN	122	66	56	20	17	10	19	0	0	2	20	17	11	6
23	OR	122	53	69	16	9	8	15	5	0	16	16	10	18	9
24	IT	111	39	72	13	10	13	2	1	0	4	24	17	11	16
25	THAN	111	45	66	5	8	18	13	1	0	23	9	10	10	14
26	AT	108	38	70	13	8	6	11	0	0	25	14	9	9	13
27	HAVE	101	69	32	15	10	17	26	1	0	1	5	6	7	13
28	HAS	98	50	48	7	14	14	14	1	0	6	4	13	9	16
29	OVER	98	55	43	6	13	19	16	1	0	9	11	8	6	9
30	TRAFFIC	96	47	49	7	10	20	8	2	0	2	3	9	18	17
31	MHZ	95	43	52	3	3	3	11	23	0	2	7	30	4	9
32	DATA	92	34	58	12	12	4	2	4	0	6	6	18	11	17
33	SIGNAL	89	53	36	11	3	6	7	26	0	3	1	6	13	13

34	NOT	83	41	42	12	9	14	6	0	0	2	13	14	9	4
35	REQUIREMENTS	79	16	63	1	5	6	3	1	0	45	5	3	7	3

Tabla 10.3 Colocados de *bandwidth*.

En la tabla se aprecian 35 de los 928 colocados que despliega el programa (Apéndice VI), registrados por orden descendente de frecuencia. Al lado de cada uno, se halla el número total de veces que co-ocurren con el nodo en las líneas de concordancia, junto a las veces que aparecen a la izquierda y a la derecha. A continuación, los totales se desglosan de acuerdo con el puesto que ocupa el colocado respecto al nodo, en una distancia de cinco palabras a la izquierda y cinco a la derecha. De esta forma, se distingue exactamente cuántas veces y a qué distancia co-ocurre el colocado con el nodo. Finalmente, el color rojo marca al colocado más frecuente y el asterisco la posición central donde se sitúa el nodo.

Los resultados obtenidos muestran la frecuencia y la posición de las palabras que co-ocurren con el nodo, pero no revelan cuáles demuestran ser colocados significativos. Para saber si nodo y colocados manifiestan algún tipo de afinidad y si la relación existente entre ellos no se debe a la casualidad, es preciso realizar una prueba estadística.

La colocación significativa en términos estadísticos se define como “*the probability of one lexical item (the node) co-occurring with another word or phrase within a specified linear distance or span being greater than might be expected from pure chance*” (Oakes, 1998:163). Existen varias medidas que cuantifican tal probabilidad, entre las cuales WordSmith permite calcular el índice de atracción mutua o *Mutual Information score* (MI). Sin embargo, el programa ofrece como máximo 10 colocados significativos y no siempre de todas las palabras. Incluso no revela los colocados de algunas de las palabras más relevantes. Como consecuencia de ello, se han transportado a una hoja de Excel los datos requeridos para realizar las operaciones pertinentes, de modo que sea posible aplicar las pruebas estadísticas más apropiadas: *t-score*, *z-score* y MI (Barnbrook, 1996). Estas medidas permiten reflejar la probabilidad y así valorar si la co-ocurrencia entre nodo y colocado es estadísticamente significativa.

La primera prueba, el índice de atracción mutua, se aplica mediante la siguiente fórmula:

$$MI = \log_2 \frac{O}{E}$$

siendo O la frecuencia observada del colocado en el entorno del nodo, es decir, el número de veces que el colocado co-ocurre con el nodo; y E la frecuencia esperada del

colocado, esto es, la frecuencia con la que se espera encontrar el colocado alrededor del nodo, calculada como se indica a continuación:

$$\text{Frecuencia esperada} = \frac{F}{F_{\text{Total}}} \times T_{\text{conc}}$$

donde F es la frecuencia absoluta del colocado en el corpus, T_{total} es el número total de *tokens* en el corpus y T_{conc} el número total de palabras acotadas en las líneas de concordancia del nodo. De este modo, se introducen los valores correspondientes para calcular la frecuencia esperada de un colocado de *bandwidth*, por ejemplo *network*. Se conoce la frecuencia absoluta de *network* (16.649), el número total de *tokens* (5.533.705) y el total de palabras acotadas (31.190). Esta última cifra deriva de multiplicar el número de líneas de concordancia obtenidas (3.119) por 10, la suma de las 5 palabras a la izquierda y a la derecha que constituyen el entorno léxico del nodo.

$$\text{Frecuencia esperada} = \frac{16.649}{5.533.705} \times 31.190 = 93,83$$

Una vez que se obtienen todos los valores necesarios, se sustituyen en la fórmula original y se halla el índice de información mutua de *network*:

$$MI = \log_2 \frac{162}{93,83} = 0,79$$

El índice obtenido señala con qué fuerza *bandwidth* atrae a *network* y si éste es un colocado significativo. La afinidad entre dos palabras será mayor cuanto más alto sea el valor conseguido en el test. El punto de corte cae en el tres, de acuerdo con Scott (1998): “*below 3.0 the linkage between node and collocate is likely to be rather tenuous*”. Por consiguiente, la atracción que *bandwidth* ejerce hacia *network* no es lo suficientemente fuerte como para que *network* constituya un colocado significativo.

La relación existente entre el nodo y colocado del ejemplo es un caso de colocación ascendente, puesto que la frecuencia absoluta de *network* es mayor que la de *bandwidth*, y este tipo de colocación no es válida para aclarar el entorno semántico del nodo. Así que, para captar aquellas palabras cuya presencia se debe a la atracción que ejerce el nodo, es preciso enfocar el análisis hacia la colocación descendente. Además, también se tendrán en cuenta las siguientes recomendaciones procedentes de la literatura

disponible (Sinclair, 1991; Barnbrook, 1996; Jackson, 1998; Scott, 1998; Nelson, 2000; Almela et al., 2005). Para identificar los colocados significativos:

1. Se computan sólo los colocados cuya frecuencia absoluta es menor que la frecuencia absoluta del nodo.
2. No se consideran las palabras funcionales dado que su presencia es tan probable que apenas constituyen una colocación significativa.
3. La frecuencia de unión entre nodo y colocado debe ser mayor que 5, con el fin de evitar la inclusión de co-ocurrencias poco relevantes.
4. La puntuación obtenida en los tests MI y Z debe ser mayor que 3.
5. La puntuación obtenida en el test T debe ser mayor que 2.

Respecto a las puntuaciones Z y T, éstas se calculan aplicando las siguientes fórmulas:

$$Z = \frac{O - E}{\sigma} \quad T = \frac{O - E}{\sqrt{O}}$$

donde O es la frecuencia observada del colocado, E la frecuencia esperada del colocado y σ la desviación estándar de la ocurrencia del colocado en el corpus entero. La desviación estándar se calcula de la siguiente manera:

$$\sigma = \sqrt{N(p(1-p))}$$

siendo N el número de palabras acotadas en las líneas de concordancia del nodo y p la probabilidad de encontrar el colocado en el entorno del nodo. Tal probabilidad se halla dividiendo la frecuencia absoluta del colocado entre el número total de *tokens* en el corpus.

Al sustituir las variables por los valores correspondientes, *network*, como colocado de *bandwidth*, obtiene las siguientes puntuaciones:

$$Z = \frac{162 - 93,59}{\sqrt{93,52}} = 7,05 \quad T = \frac{162 - 93,59}{\sqrt{162}} = 5,35$$

Los colocados de *bandwidth* que cumplen con todos los requisitos precedentes ascienden a 82, los cuales se muestran a continuación junto a la puntuación obtenida en los diferentes tests realizados:

Nº	Colocados	Z	T	MI	Nº	Colocados	Z	T	MI
1	ADAPTATION	26,81	5,80	4,42	42	LI	7,16	2,21	3,38
2	AFFECTS	8,49	2,57	3,45	43	LIMITATIONS	11,66	4,03	3,06
3	AGGREGATE	21,05	4,83	4,25	44	LVCM	12,94	2,17	5,15
4	ALLOCATE	20,42	4,18	4,58	45	MAXIMIZE	10,50	2,92	3,69
5	ALLOCATED	35,44	7,06	4,66	46	MAXIMIZING	7,66	2,07	3,77
6	ALLOCATION	35,98	7,91	4,37	47	MB	8,76	2,71	3,38
7	ALLOCATIONS	8,69	2,10	4,09	48	MBPS	17,05	5,07	3,50
8	ALPHA	8,43	2,43	3,59	49	MHZ	42,85	9,29	4,41
9	AMOUNT	24,74	7,67	3,38	50	MILE	6,99	2,21	3,33
10	BAAS	13,85	2,18	5,33	51	MINMAX	21,86	2,78	5,95
11	BI	11,35	3,07	3,77	52	MULTIMEDIA	14,08	4,83	3,09
12	BOTTLENECK	8,37	2,56	3,42	53	NARROW	17,79	4,49	3,97
13	BT	9,15	2,30	3,99	54	OCCUPIED	46,76	5,74	6,05
14	BW	23,31	3,25	5,68	55	OCTETS	9,32	2,61	3,68
15	CBWFQ	14,56	2,19	5,47	56	PROVISIONING	9,64	3,10	3,27
16	COMMODITY	10,87	2,66	4,06	57	RAW	6,55	2,32	3,00
17	CONSERVE	14,44	2,38	5,20	58	RESERVABLE	43,32	2,82	7,89
18	CORRELATOR	7,59	2,07	3,75	59	RESERVE	13,92	2,87	4,55
19	DECOUPLING	6,88	2,04	3,51	60	RESERVED	24,58	5,23	4,47
20	DEGRADED	12,02	2,53	4,50	61	RESIDUAL	6,54	2,18	3,17
21	DEMAND	23,38	7,04	3,46	62	SAVES	8,02	2,08	3,89
22	DEMANDS	10,83	3,57	3,20	63	SCARCE	11,41	2,35	4,56
23	DOWNSTREAM	9,45	3,29	3,04	64	SCHEDULER	12,74	3,36	3,85
24	EFFICIENCY	18,56	6,36	3,09	65	SHANNON	15,76	2,90	4,89
25	EFFICIENT	23,26	7,63	3,22	66	SHAPING	8,83	2,72	3,40
26	EXCESS	25,24	5,08	4,63	67	SNRS	14,66	2,56	5,03
27	FAIR	7,45	2,51	3,14	68	TB	10,54	2,33	4,36
28	FAIRNESS	17,02	4,01	4,17	69	THZ	21,97	2,21	6,62
29	FDM	10,06	2,14	4,47	70	TLV	20,42	3,37	5,20
30	FI	7,66	2,64	3,07	71	TRADING	12,71	2,99	4,18
31	GBIT	9,56	2,31	4,10	72	TRANSPONDER	9,92	2,32	4,20
32	GHZ	15,23	5,22	3,09	73	UNIFORMLY	8,64	2,28	3,85
33	GUARANTEED	12,91	3,86	3,48	74	UNITY	6,18	2,15	3,04
34	GUARANTEES	13,55	3,80	3,67	75	UNRESERVED	45,11	2,99	7,83
35	HAUL	9,55	2,47	3,90	76	UNUSED	19,29	3,61	4,84
36	HOGGING	33,15	2,23	7,79	77	UPSTREAM	13,50	3,98	3,52
37	HUGE	11,26	3,50	3,37	78	USAGE	15,27	5,22	3,09
38	INTENSIVE	17,48	4,12	4,17	79	UTILIZATION	30,73	6,15	4,64
39	ISPS	8,26	2,56	3,38	80	UTILIZE	11,95	3,21	3,79
40	KBPS	21,55	5,01	4,21	81	WIDER	19,43	4,35	4,32
41	KHZ	27,10	5,43	4,64	82	WS	9,47	2,30	4,08

Tabla 10.4 Colocados significativos de *bandwidth*.

Los resultados obtenidos muestran la selección léxica que la presencia de *bandwidth* condiciona con mayor o menor grado de cohesión. El léxico captado se

encuentra íntimamente relacionado con el concepto de *bandwidth* y contribuye a la vez a construir su significado.

En la definición de *bandwidth* plasmada en el diccionario especializado Webster disponible en internet (www.websters-online-dictionary.com), se pueden observar las conexiones entre los componentes semánticos de *bandwidth* y los colocados significativos que atrae:

Bandwidth (Noun): *A data transmission rate; the maximum amount of information (bits/second) that can be transmitted along a channel.*

Specialty Definition:

Bandwidth refers to the width, usually measured in Hertz, of a frequency band. It can also be used to describe a signal, in which case the meaning is the width of the smallest frequency band within which the signal can fit. Bandwidth is related to the amount of information that can flow through a channel through the Nyquist-Shannon sampling theorem.

The ***Bandwidth*** of an electronic filter is the part of the filter's frequency response that lies within 3 dB of its peak.

The term ***Bandwidth*** is also used, informally, and by extension from the above, to mean the amount of data that can be transferred through a connection in a given time period. Bandwidth is normally based on the frequencies used and the spectral spread of the information carried on the frequency.

El concepto de colocación al que se ha hecho referencia, alude a la posible atracción existente entre palabras, pero respecto a una forma individual. Sin embargo, las formas individuales pueden tener usos diferentes y formar parte de una unidad mayor de significado o de una combinación de palabras recurrente. Además, en un sublenguaje de especialización, es posible que estas agrupaciones estén provistas de significado específico: “*Very frequent words in specialized corpora in fact often tend to aggregate in recurrent chunks to form more specialised meanings*” (Gavioli, 2005:79). Tales combinaciones de palabras son identificables mediante la observación del nodo en su contexto, examinando la palabra inmediatamente anterior y posterior, o bien por medio de los clusters.

El hecho de que la palabra *bandwidth* aparezca 3.113 veces en el corpus, conlleva afrontar el mismo número de líneas de concordancia cuando se requiere estudiar el contexto, una cifra muy complicada de manejar (Apéndice VI). Por lo tanto, es necesario reducir el volumen de datos y extraer una muestra. En este caso, se recurre a los colocados de mayor frecuencia de unión que preceden y suceden al nodo en primera posición.

a) Colocado + *bandwidth*.

Los colocados seleccionados aparecen en posición pre-nominal con una frecuencia mínima de 8, siendo 119 la frecuencia del colocado más recurrente. El total de colocados asciende a 50, entre los cuales los adjetivos superan al resto de categorías, seguidos de los sustantivos, adverbios y verbos. Se añaden las siglas como una forma compleja peculiar que precede al nodo.

- Adjetivos, total 21: *high* 119/*higher* 55, *available* 86, *low* 33/*lower* 13, *maximum* 23, *total* 20, *same* 15, *wide* 15/*wider* 11, *greater* 14, *additional* 13, *narrow* 13, *optical* 12, *large* 11/*larger* 8, *aggregate* 10, *different* 10, *full* 10, *sufficient* 8, *variable* 8.
- Sustantivos, total 15: *network* 47, *signal* 26, *excess* 25, *gain* 24, *link* 24, *memory* 24, *information* 21, *loop* 17, *impedance* 14, *channel* 12, *backbone* 10, *communication* 10, *input* 10, *transition* 10, *modulation* 8.
- Adverbios, total 5: *more* 58, *enough* 16, *much* 14, *less* 11, *upstream* 11.
- Formas personales de verbos, total 1: *limit* 8.
- Participios verbales que actúan como adjetivos, total 5: *occupied* 30, *allocated* 21, *limited* 19, *reduced* 15, *reserved* 13.
- Siglas, total 3: *CPU* 9, *ghz* 9, *WAN* 9.

b) *Bandwidth* + colocado.

Los colocados extraídos en posición post-nominal presentan una frecuencia mínima de 8 y una máxima de 53. El total de colocados en esta posición se reduce a 23, con un marcado predominio de sustantivos.

- Sustantivos, total 17: *allocation* 53, *requirements* 45/*requirement* 10, *efficiency* 39, *utilization* 27, *management* 25, *adaptation* 23, *usage* 21, *values* 18, *product* 15, *needs* 15, *capacity* 11, *services* 10, *availability* 9, *demands* 9, *resources* 9, *consumption* 8.
- Adjetivos, total 2: *efficient* 42, *intensive* 17.
- Participios verbales que actúan como adjetivos, total 2: *required* 16, *scheduling* 10.
- Formas personales de verbos, total 1: *used* 14.
- Participios verbales que actúan como sustantivos, total 1: *sharing* 8.

Una vez que se han identificado los colocados contiguos más frecuentes, se exploran las líneas de concordancia correspondientes, con el fin de inferir por el contexto si tales agrupaciones adquieren un significado especializado. A modo de ejemplo, el Cuadro 10.1 presenta todas las líneas de concordancia donde aparece la combinación *memory bandwidth*, y que evidencian el uso y significado especializado adquirido.

...cture becomes commercially available. Whilst **memory bandwidth** is the key to speech recognition, let's not fo
t but it is impractical due to limitations in the **memory bandwidth** [7]. Figure 3 shows an example of the tran
e IXP1200). In this case we observed that the **memory bandwidth** remains constant for MPLS, whereas *IntSe*
can be explained by the availability of higher **memory bandwidth** in the four-processor machine when compar
use of multiple processors taxes the available **memory bandwidth** and can lead to lengthy stall times if a cac
for further investigation. Considering today's **memory bandwidth** of 80Gbit/s, an architecture with 16 = = 1 n
E of the increase of threads. As result of this, **memory bandwidth** in *IntServ* grows proportionally with the num
ure enables us to evaluate the influence of the **memory bandwidth** on communication and computation perfor
odeling parameters in these expressions are: **Memory bandwidth (BWm)**: Their units are bits per second (bp
nough internal registers to compensate for its **memory bandwidth** shortfalls. The PowerPC, on the other hand,
ds to equation (4). However, the model for the **Memory Bandwidth** takes into account the hashing collisions a
Serv6 present an increase on number of flows. **Memory bandwidth** in *IntServ6* has a better behaviour than *IntS*
architectures provided improved processor-to-**memory bandwidth** by introducing separate data and instruction
ndwidth of MPLS. In addition, the increase of **memory bandwidth** in *IntServ6* is consequence of the use of 12
due to the use of the Tagging, for this reason **memory bandwidth** for these two architectures stays constant.
. hash table readings. 4.2. Relation between **memory bandwidth (BWm)** and flows number. These results ar
are observed in figures 4 and 5. We simulated **memory bandwidth** under the following conditions: - Hash tabl
(m) is constant in 6400. Figure 3. Normalized **memory bandwidth** vs. hash table readings. 4.2. Relation betw
idth (BWm) and hash table readings (k). The **memory bandwidth** based on k was simulated, maintaining the
ed in the next section. 4.1. Relation between **memory bandwidth (BWm)** and hash table readings (k). The m
be very effective in increasing the processor-to-**memory bandwidth** and hence provide increased system perfor
mand for bandwidth far exceeds the growth in **memory bandwidth**, making it inevitable that switches will maint
h recognition is pattern matching, which takes **memory bandwidth**. The main benefit of a dsp is its ability to pe
2 bits. Figure 3 shows variation of normalized **memory bandwidth**. The variation of these parameters affect In

Cuadro 10.1 Concordancias de *memory bandwidth*.

Extendiendo el procedimiento seguido, se obtienen las siguientes combinaciones, donde se incluyen las variantes en las que los colocados cambian de categoría, pero no de significado. Los colocados que se identificaron como significativos se resaltan en negrita, con la intención de reflejar su distribución e influencia en el entorno más próximo del nodo:

a) Colocado + *bandwidth*:

<i>aggregate/aggregated bandwidth</i>	<i>information bandwidth</i>	<i>optical bandwidth</i>
<i>backbone bandwidth</i>	<i>link bandwidth</i>	<i>signal bandwidth</i>
<i>channel bandwidth</i>	<i>loop bandwidth</i>	<i>total bandwidth</i>
<i>CPU bandwidth</i>	<i>memory bandwidth</i>	<i>transition bandwidth</i>
<i>excess bandwidth</i>	<i>modulation bandwidth</i>	<i>upstream bandwidth</i>
<i>gain bandwidth</i>	<i>network bandwidth</i>	<i>WAN bandwidth</i>
<i>impedance bandwidth</i>	<i>occupied bandwidth</i>	

b) *Bandwidth* + colocado:

bandwidth adaptation

bandwidth scheduling/scheduler

bandwidth allocation/allocated/allocations

bandwidth sharing/share

bandwidth consumption/consumed

Bandwidth value/values

bandwidth management

El siguiente apartado está dedicado al estudio de combinaciones de más de dos unidades, con el fin de completar el perfil léxico del registro especializado y reflejar el uso idiomático del lenguaje.

10.2.4 Agrupaciones léxicas o clusters.

Los clusters son agrupaciones léxicas que se encuentran repetidamente en compañía. A diferencia de las colocaciones, las unidades establecen una relación mucho más fuerte, puesto que son repeticiones exactas de una misma secuencia. Tales secuencias están formadas por una extensión determinada, es decir, existen clusters de dos, tres, cuatro o más palabras y que, en ocasiones, es posible hallar en una estructura embebida. Por lo tanto, en el presente análisis, se parte de los clusters de dos palabras identificados previamente, con el fin de comprobar si éstos pertenecen a su vez a unidades léxicas mayores. Paralelamente, se examinan las líneas de concordancia de los clusters, en su variedad de extensiones, en búsqueda de otros patrones de colocación.

En el Apéndice VI se hallan los listados de clusters, de dos a seis unidades, que se producen a lo largo de todas las líneas de concordancia de *bandwidth* con una frecuencia mínima de tres, distribuidos de la siguiente forma:

Combinaciones	Volumen	Frecuencia máxima
2-cluster	1548	442
3-cluster	681	94
4-cluster	204	40
5-cluster	62	6
6-cluster	17	4

Cuadro 10.2 Clusters de diferentes unidades de *bandwidth*.

El hecho de que los clusters de dos palabras sean los más numerosos se debe a las combinaciones de palabras funcionales y nocionales, y de funcionales entre sí. No obstante, sólo se tendrán en cuenta las agrupaciones de palabras de contenido, donde además, uno de los componentes sea el nodo. De esta forma, el listado de clusters de dos unidades se reduce a 264.

El siguiente paso consiste en examinar los clusters de extensión superior que contienen las 28 combinaciones identificadas anteriormente como especializadas. Por

un lado, el patrón más recurrente contempla las siguientes posibilidades, producto del sistema gramatical del lenguaje:

Colocado + <i>bandwidth</i>	
3-clusters	Muestras
<i>the + colocado + bandwidth</i> <i>a/an + colocado + bandwidth</i> <i>colocado + bandwidth+ of</i> <i>colocado + bandwidth + and</i> <i>colocado + bandwidth+ is</i> <i>colocado + bandwidth + in</i> <i>in + colocado + bandwidth</i> <i>and + colocado + bandwidth</i>	<i>the impedance bandwidth</i> <i>an aggregate bandwidth</i> <i>modulation bandwidth of</i> <i>excess bandwidth and</i> <i>information bandwidth is</i> <i>signal bandwidth in</i> <i>in network bandwidth</i> <i>and transition bandwidth</i>
4-clusters	Muestras
<i>a/an + colocado + bandwidth+ of</i> <i>in + the + colocado + bandwidth</i> <i>colocado + bandwidth + of + the</i> <i>than + the + colocado + bandwidth</i>	<i>an aggregate bandwidth of</i> <i>in the information bandwidth</i> <i>total bandwidth of the</i> <i>than the information bandwidth</i>

Por otro lado, en la progresión de los clusters, se observa la inclusión de algunas de estas combinaciones especializadas en unidades más amplias:

<p><i>bandwidth allocation</i> <i>bandwidth allocation policy</i> <i>bandwidth allocation policies</i> <i>minmax bandwidth allocation</i></p> <p><i>bandwidth adaptation</i> <i>bandwidth adaptation algorithms</i> <i>bandwidth adaptation algorithms BAAS</i> <i>continuous bandwidth adaptation</i> <i>discrete bandwidth adaptation</i></p> <p><i>bandwidth management</i> <i>BPM bandwidth management</i> <i>bandwidth management agent</i></p> <p><i>bandwidth scheduling</i> <i>prediction and bandwidth scheduling</i> <i>traffic prediction and bandwidth scheduling</i></p>	<p><i>bandwidth values</i> <i>possible bandwidth values</i> <i>bandwidth values of adaptive multimedia</i></p> <p><i>loop bandwidth</i> <i>open loop bandwidth</i> <i>closed loop bandwidth</i></p> <p><i>gain bandwidth</i> <i>unity gain bandwidth</i> <i>gain bandwidth product</i></p> <p><i>occupied bandwidth</i> <i>occupied bandwidth symbol</i> <i>occupied bandwidth symbol rate</i></p>
--	--

Cuando se detectaron en el apartado anterior los colocados significativos en posición pre-nominal, se señaló la notoria presencia de adjetivos, destacando aquellos que expresan cantidad o medida: *high 119/higher 55, low 33/lower 13, maximum 23, wide 15/wider 11, greater 14, additional 13, narrow 13, large 11/larger 8, full 10, sufficient 8*. Éstos aparecen de nuevo formando clusters de dos unidades, junto a otros premodificadores que expresan igualmente cantidad o medida:

2-cluster + frecuencia					
<i>enough bandwidth</i>	17	<i>large bandwidth</i>	10	<i>more bandwidth</i>	60
<i>excess bandwidth</i>	25	<i>larger bandwidth</i>	7	<i>much bandwidth</i>	14
<i>extra bandwidth</i>	7	<i>less bandwidth</i>	11	<i>narrow bandwidth</i>	13
<i>full bandwidth</i>	10	<i>little bandwidth</i>	3	<i>reduced bandwidth</i>	17
<i>greater bandwidth</i>	14	<i>low bandwidth</i>	21	<i>remaining bandwidth</i>	10
<i>high bandwidth</i>	68	<i>low-bandwidth</i>	12	<i>small bandwidth</i>	4
<i>high-bandwidth</i>	52	<i>lower bandwidth</i>	10	<i>sufficient bandwidth</i>	8
<i>higher bandwidth</i>	30	<i>lower-bandwidth</i>	3	<i>smaller bandwidth</i>	4
<i>higher-bandwidth</i>	5	<i>lowest bandwidth</i>	3	<i>wide bandwidth</i>	11
<i>highest bandwidth</i>	3	<i>maximum bandwidth</i>	25	<i>wide-bandwidth</i>	4
<i>huge bandwidth</i>	7	<i>minimum bandwidth</i>	8	<i>wider bandwidth</i>	12

El cluster más recurrente formado por un adjetivo está constituido por *high+bandwidth* (165), seguido de *low+bandwidth* (37) y *wide+bandwidth* (27), los cuales se hallan en agrupaciones léxicas de mayor extensión. De nuevo *high bandwidth* es la combinación más prolífica, aunque sólo interviene en combinaciones de tres unidades como máximo, mientras que *low bandwidth* es menos productiva, pero participa en clusters de cuatro y cinco unidades. Por su parte, *wide bandwidth*, presenta una única combinación de tres unidades, insertadas a su vez en un cluster inmediatamente superior. Se incluyen también los clusters de *narrow bandwidth*, por ser el opuesto de *wide bandwidth*, con el fin de observar si se producen patrones léxicos comunes entre los pares *high-low* y/o *wide-narrow*.

3-cluster + frecuencia			
<i>a high bandwidth</i>	6	<i>extremely high bandwidth</i>	3
<i>a high-bandwidth</i>	5	<i>for high-bandwidth</i>	3
<i>a higher bandwidth</i>	5	<i>have high bandwidth</i>	3
<i>high bandwidth efficiency</i>	5	<i>high bandwidth (n°)</i>	3
<i>for high bandwidth</i>	4	<i>high bandwidth access</i>	3
<i>high frequencies have</i>	4	<i>high bandwidth applications</i>	3
<i>high-bandwidth applications</i>	4	<i>high bandwidth requirements</i>	3
<i>higher bandwidth and</i>	4	<i>high-bandwidth data</i>	3
<i>of high-bandwidth</i>	4	<i>the highest bandwidth</i>	3
<i>providing high bandwidth</i>	4	<i>their high bandwidth</i>	3
<i>the high bandwidth</i>	4	<i>very high bandwidth</i>	3
<i>the higher bandwidth</i>	4	<i>high-bandwidth services</i>	3
<i>a higher-bandwidth</i>	3	<i>very high bandwidth</i>	3

4-cluster + frecuencia		3/4-cluster + frecuencia			
<i>a very low bandwidth</i>	3	<i>very wide bandwidth</i>	4	<i>a narrow bandwidth</i>	4
<i>low bandwidth redundancy to</i>	3	<i>a very wide bandwidth</i>	4	<i>narrow bandwidth of</i>	3
<i>very low bandwidth redundancy</i>	3				
5-cluster + frecuencia					
<i>a very low bandwidth redundancy</i>	3				
<i>requires a very low bandwidth</i>	3				
<i>very low bandwidth redundancy to</i>	3				

De las líneas de concordancia de *bandwidth*, se extraen aquellas donde intervienen los clusters de dos unidades seleccionados y se examinan las palabras situadas a su derecha, de tal modo que se evidencian los siguientes comportamientos:

High bandwidth + (premodificador) + sustantivo: *(broadcast oriented) applications, (transmission) capabilities, circuits, (wireless/dedicated) connection(s), efficiency, (IP) technologies, links, pipe, requirements, (data transmission) services, (digital internet) access, (memory) interface, users, (transport/3G) networks, amplifiers, (communication) channels, data, fiber, flows, I/O, performance, traffic, management, transport, capability, bus, (data) path, switch.*

Low bandwidth + (premodificador) + sustantivo: *availability, continuous data services, (data) links, redundancy, traffic, (microwave radio) applications, communication infrastructures, (digitized) information, network, video, WAN connections, requirements, utilization, vocoders.*

Wide bandwidth + sustantivo: *amplifiers, communications, signal(s), dual-polarization designs, matching technique, microwave, transmission.*

Narrow bandwidth + sustantivo: *link, connections, source.*

Entre ellos se manifiestan patrones comunes, especialmente entre *high* y *low*, de forma que las siguientes combinaciones pertenecen al uso idiomático del lenguaje de la ingeniería de telecomunicación:

<i>High / low bandwidth applications</i>	<i>High / low bandwidth traffic</i>
<i>High / low bandwidth network</i>	<i>High / wide bandwidth amplifier</i>
<i>High / low bandwidth services</i>	<i>High / low / narrow bandwidth connection</i>
<i>High / low bandwidth requirements</i>	<i>High / low / narrow bandwidth links</i>

Finalmente, se muestran los colocados contiguos de las combinaciones especializadas detectadas en las líneas de concordancia. De nuevo, se marca en negrita la presencia de los colocados significativos, con el fin de resaltar su proximidad y/o inclusión en las combinaciones especializadas. El hecho de que los colocados significativos se encuentren tan cerca de las combinaciones, y que incluso a veces lleguen a constituir clusters de mayores unidades, les confiere una gran importancia. Por lo tanto, estas combinaciones y especialmente las que contienen colocados significativos, deberían aparecer en los libros de texto destinados a la enseñanza del inglés profesional y académico de las telecomunicaciones.

pre +	COMBINACIÓN	+ post
<i>double, greater</i>	<i>aggregate(d) bandwidth</i>	<i>allotment, availability, obtained, results, usage</i>
<i>PBX, currently, total, per-VC, system</i>	<i>allocate(d) bandwidth</i>	<i>based, during, resource</i>
<i>maximum, probed, increases, actual, magnifying, fluctuating, large, average, affects, total, continuously</i>	<i>available bandwidth</i>	<i>depends, permissible, used</i>
<i>managing, wireless,</i>	<i>backbone bandwidth</i>	<i>needs</i>

<i>aggregate, shared</i>		
<i>kHz, passband, band, RF</i>	<i>channel bandwidth</i>	<i>requirement, used</i>
<i>allocated, total</i>	<i>CPU bandwidth</i>	<i>counter, memory, partitioning</i>
<i>large, smaller, distribute, versus, signal, fractional, available, unused</i>	<i>excess bandwidth</i>	<i>fairly, helps, provided, present</i>
<i>unity, constant, demands, infinite, achieved, amplifier, Op-Amp, high, lower, circuit's, closed-loop</i>	<i>gain bandwidth</i>	<i>product(s), GBW</i>
<i>wider, good, maximum, comparable, narrow, input</i>	<i>impedance bandwidth</i>	<i>VSWR, compromises</i>
<i>high, channel's, limited</i>	<i>information bandwidth</i>	<i>decreases, occupies</i>
<i>total, wireless, free, cable, available, valuable</i>	<i>link bandwidth</i>	<i>corresponding, margin, status, delay</i>
<i>optimal, open, closed, carrier, effective</i>	<i>loop bandwidth</i>	<i>design, helps, coincide</i>
<i>available, processor-to-, reason, higher, today's, normalized, takes, simulated</i>	<i>memory bandwidth</i>	<i>BWm, remains, shortfalls, takes</i>
<i>narrow, intensity</i>	<i>modulation bandwidth</i>	
<i>increased, conserve, available, providing, supplying, greater, ensuring, least, reserve, aggregate, regarding, saves</i>	<i>network bandwidth</i>	<i>allocation, available, capacity, delay, increases, requirements, resources, usage</i>
<i>wide, signals, typical, ideal</i>	<i>occupied bandwidth</i>	<i>BW, numbers</i>
<i>expand, all-, 3dB</i>	<i>optical bandwidth</i>	<i>transiting</i>
<i>total, optical, ST, transmitted, constant, input, higher filtered, original, desired, reduced,</i>	<i>signal bandwidth</i>	<i>capability</i>
<i>increased, available, overall</i>	<i>transition bandwidth</i>	
<i>transmits, limited, hostile, usable, Kbps</i>	<i>upstream bandwidth</i>	<i>allocation, efficiency, limited</i>
<i>optimizes, understanding, minimal</i>	<i>WAN bandwidth</i>	<i>delivery, needs, reduction, usage</i>
<i>continuous, discrete, optimal, propose(d), two, near-optimal</i>	<i>bandwidth adaptation</i>	<i>case, algorithms, framework</i>
<i>adjust, applying, best, demand, different, dynamic, fixed, flexible, improve, MinMax, network, new, OPS, optimal, peak, previous, take, transponder, upstream, user, certain, inaccurate, existing</i>	<i>bandwidth allocation(s)</i>	<i>model, based, message, algorithm, mechanisms, policy, policies, accuracy, request, flexibility, performance, maps, possible</i>
<i>caused, LAN; reduce(d)</i>	<i>bandwidth consumed</i>	<i>problems</i>
<i>good, high, increase, maximizing, theoretical, upstream, poor, provides</i>	<i>bandwidth efficiency</i>	<i>describes, limitations, limits, technique(s), constant, type, operating, fading, modulation</i>
<i>market-based, comprehensive, dynamic, includes, QoS, adaptive, application-aware</i>	<i>bandwidth management</i>	<i>capability, role, tools, optimization, support, becomes, features, agent</i>
<i>former, future, MinMax</i>	<i>bandwidth scheduling</i>	<i>technique, problem, algorithm(s)</i>
<i>entire, fair, fine-grained, unavoidable, better</i>	<i>bandwidth sharing</i>	
<i>discrete, exact, maximum, target, previous, largest, possible</i>	<i>bandwidth value(s)</i>	

Tabla 10.5 Colocados contiguos de las combinaciones especializadas.

Los tres apartados siguientes contienen el análisis detallado de las formas más significativas, elegidas como representantes entre los miembros de sus familias. Primero se analizan las formas técnicas, después las formas académicas y, por último, las formas generales. Seguidamente, se presentan las conclusiones obtenidas de tales análisis.

10.3 Análisis detallado de formas técnicas.

A continuación se exponen las fichas correspondientes al análisis realizado de una muestra de formas representantes de familias identificadas como especializadas: *protocol*, *wireless*, *chip*, *multicast*, *cache*, *throughput*, *latency*, *impedance*, *bluetooth*, *firewall* y *crosstalk*. En el Apéndice VI, se hallan las fichas de todas las formas analizadas, junto a los listados completos de los colocados, colocados significativos, líneas de concordancia y clusters. Igualmente, el Apéndice VI incluye las fichas y datos correspondientes a: *cosine*, *diffraction*, *netlist*, *satellite*, *applet*, *microstrip*, *dialog*, *unicast* y *timeslot*.

1. Familia de *protocol*. Forma analizada: PROTOCOL

1. FRECUENCIA.

KEYWORD	Frec. Telec	Frec. Lacell	Ratio	Término, Chung	KEYNESS	P valor
7. PROTOCOL	4.742	139	129,564867	ESPECIFICO	13.677,70	0
Miembros relacionados:	F Telec	F Gene	Ratio	Término	Keyness	P valor
PROTOCOLS	1.996	52	145,779917	ESPECIFICO	5.800,40	0
MULTIPROTOCOL	73	0	inf/esp	inf/esp	229	0
Familia técnica: Sí, todos los miembros son valorados como específicos.						

2. DISTRIBUCIÓN.

Distribución por áreas	Distribución Keyness	Keykeyword	Nº de textos	Porcentaje
9	3,082	PROTOCOL	279	16,87 %
<i>Dispersion plot</i>				
Area	Words	Hits	Per 1,000	
082 Esp.Tele	997.694	2613	2.62	
3 Telematics	1.204.957	1268	1.05	
081 Esp.Sing	867.140	465	0.54	
7 Systems	307.660	75	0.24	
4 Signal proc.	580.874	133	0.23	
6 Business	373.038	70	0.19	
2 Ar. Comp	329.602	44	0.13	
1 Electronics	722.798	71	0.10	
5 Materials	101.237	2	0.02	

3. COLOCACIÓN.

3.1 Colocados.

N	Collocates	TOTAL	LEFT	RIGHT	L5	L4	L3	L2	L1	*	R1	R2	R3	R4	R5
1	THE	3636	2306	1330	319	414	460	659	454	0	6	398	356	285	285
2	A	1324	858	466	125	144	251	168	170	0	5	117	142	111	91
3	IS	1130	405	725	135	146	72	42	10	0	228	183	126	86	102
4	OF	1047	712	335	98	130	214	223	47	0	10	26	56	110	133
5	AND	1013	425	588	109	109	94	65	48	0	91	197	93	106	101
6	TO	902	371	531	94	92	90	86	9	0	77	94	112	133	115
7	IN	692	284	408	62	77	77	56	12	0	68	92	83	85	80
8	FOR	565	179	386	34	46	42	46	11	0	111	76	65	70	64
9	ROUTING	538	451	87	21	22	17	51	340	0	1	13	24	20	29
10	THAT	442	127	315	26	30	34	32	5	0	117	71	35	53	39
11	NETWORK	377	258	119	32	33	33	95	65	0	3	13	29	34	40
12	INTERNET	356	269	87	22	14	45	15	173	0	15	6	28	28	10
13	LAYER	353	258	95	38	24	35	34	127	0	10	5	10	22	48
14	AS	346	179	167	41	56	52	21	9	0	12	49	40	46	20
15	THIS	309	144	165	27	24	17	9	67	0	0	61	34	41	29
16	AN	304	177	127	30	41	56	50	0	0	1	29	42	22	33
17	CONTROL	277	214	63	11	11	8	72	112	0	0	4	17	19	23
18	IP	272	112	160	16	21	23	29	23	0	61	14	37	24	24
19	BY	265	142	123	25	31	56	22	8	0	5	39	26	27	26

20	ARE	250	86	164	33	32	11	8	2	0	23	39	37	31	34
21	ON	238	115	123	33	21	27	27	7	0	7	36	24	25	31
22	WITH	233	107	126	33	24	28	19	3	0	23	32	23	19	29
23	USED	213	43	170	12	16	7	7	1	0	46	39	44	23	18
24	DATA	207	90	117	21	22	20	23	4	0	35	13	18	26	25
25	ACCESS	205	152	53	12	8	12	9	111	0	4	5	8	14	22
26	OSPF	201	135	66	13	9	12	21	80	0	3	19	15	18	11
27	TCP	197	84	113	23	18	14	5	24	0	38	13	36	13	13
28	BE	191	56	135	20	23	10	2	1	0	1	48	33	29	24
29	TRANSPORT	190	148	42	2	3	5	31	107	0	0	5	3	21	13
30	IT	181	54	127	19	16	14	5	0	0	2	53	29	15	28
31	WHICH	173	42	131	9	11	11	6	5	0	19	68	22	12	10
32	INFORMATION	165	99	66	20	11	7	5	56	0	6	6	11	26	17
33	OR	164	71	93	17	25	20	6	3	0	17	29	12	22	13
34	LINK	158	103	55	10	7	28	51	7	0	0	11	7	23	14
35	GATEWAY	151	119	32	4	2	1	37	75	0	0	1	8	18	5

3.2 Colocados significativos.

Nº	Colocados	Z	T	MI	Nº	Colocados	Z	T	MI
1	INTERSOCKET	37,26	2,44	7,87	87	RTP	18,41	4,14	4,31
2	PROTOCOL'S	41,61	3,71	6,97	88	TRANSPORTING	11,83	2,68	4,28
3	SUITE'S	28,34	2,62	6,87	89	LMP	9,35	2,12	4,28
4	CFSMS	23,95	2,22	6,87	90	ARP	22,70	5,19	4,26
5	ATP	27,04	2,80	6,54	91	IGRP	23,02	5,28	4,25
6	EGP	41,44	4,31	6,53	92	TRANSPORT	56,15	13,04	4,21
7	TOKENTALK	24,93	2,62	6,50	93	ENCAPSULATION	19,07	4,44	4,21
8	ETHERTALK	24,93	2,62	6,50	94	CML	12,18	2,84	4,20
9	HAVI	29,63	3,13	6,49	95	TFTP	11,44	2,67	4,19
10	LOCALTALK	26,63	2,96	6,34	96	OSPF	57,31	13,40	4,19
11	PAP	22,49	2,61	6,21	97	VOCABULARY	13,32	3,13	4,18
12	RTCP	18,83	2,21	6,19	98	VOTING	13,76	3,27	4,15
13	CFSM	23,64	2,79	6,17	99	ROUTING	92,14	21,89	4,15
14	ELECTS	20,20	2,41	6,13	101	MULTILINK	8,87	2,11	4,14
15	RTMP	25,82	3,12	6,10	102	IPX	19,84	4,72	4,14
16	NBP	21,51	2,61	6,09	103	RFC	28,34	6,80	4,12
17	DMI	17,73	2,20	6,02	104	HELLO	32,98	7,94	4,11
18	INTERIOR	62,19	7,81	5,99	105	EIGRP	12,34	2,98	4,10
19	NSP	16,79	2,20	5,87	106	ELECTION	15,59	3,77	4,10
20	APPLETALK	75,63	10,02	5,83	107	LANE	13,48	3,26	4,09
21	SUITES	45,36	6,05	5,81	108	OSI	37,50	9,08	4,09
22	TP	26,05	3,54	5,76	109	COHERENCE	12,82	3,12	4,08
23	DAP	26,05	3,54	5,76	110	SNMP	29,14	7,10	4,07
24	INITIATION	32,10	4,49	5,67	111	PDUS	10,75	2,66	4,03
25	LDP	40,08	5,63	5,66	112	MICROPAYMENT	20,99	5,22	4,01
26	LIGHTWEIGHT	47,72	6,86	5,60	113	BGP	27,20	6,82	3,99
27	IDRP	15,83	2,39	5,45	114	DHCP	18,26	4,59	3,98
28	EAP	22,39	3,38	5,45	115	CR	10,54	2,65	3,98
29	DDP	26,45	4,03	5,43	116	LDAP	29,03	7,31	3,98
30	DATAGRAM	51,75	7,93	5,41	117	FACTO	9,04	2,29	3,96
31	SEQUENCED	15,34	2,39	5,36	118	RESERVATION	23,40	5,98	3,93
32	RTSP	15,34	2,39	5,36	119	SPP	8,07	2,09	3,90
33	PIM	30,82	4,87	5,32	120	SMDS	11,80	3,09	3,87
34	EBGP	13,59	2,18	5,28	121	SDLC	9,41	2,46	3,87
35	INTERDOMAIN	18,57	3,08	5,19	122	MULTIPROTOCOL	10,59	2,79	3,85

36	STACKS	25,56	4,24	5,18	123	SNAP	10,42	2,79	3,81
37	FILING	16,48	2,75	5,17	124	NOVELL	13,46	3,60	3,81
38	EXTERIOR	27,50	4,66	5,12	125	ADVERTISEMENT	12,42	3,34	3,79
39	HYPER	12,72	2,17	5,10	126	XEROX	8,41	2,27	3,78
40	SUITE	64,57	11,03	5,10	127	NTP	12,79	3,47	3,77
41	XNS	19,65	3,36	5,09	128	IMPLEMENTS	17,70	4,81	3,76
42	CERTIFICATIONS	14,98	2,57	5,09	129	PDU	14,79	4,03	3,75
43	AEP	13,72	2,38	5,06	130	TCP	47,59	12,99	3,75
44	ROUTABLE	16,66	2,91	5,04	131	ETSI	9,32	2,61	3,68
45	CLNP	13,38	2,37	4,99	132	EXCHANGES	13,50	3,80	3,66
46	SS	23,00	4,11	4,97	133	SESSION	31,17	8,78	3,66
47	HYPERTEXT	21,41	3,87	4,94	134	MEMO	10,31	2,91	3,65
48	IGMP	22,03	3,99	4,93	135	DISCOVERY	17,23	4,87	3,65
49	SSP	13,07	2,37	4,93	136	SAP	10,24	2,91	3,63
50	TEA	15,01	2,73	4,91	137	AUTHENTICATED	11,04	3,18	3,59
51	NETWARE	28,35	5,20	4,89	138	ANNOUNCEMENT	7,07	2,05	3,57
52	PNNI	15,65	2,90	4,87	139	SMRP	7,73	2,24	3,57
53	GNUTELLA	14,75	2,73	4,87	140	FTP	19,11	5,63	3,53
54	IGP	22,58	4,21	4,85	141	STANDBY	7,50	2,23	3,50
55	CONNECTIONLESS	26,33	4,92	4,84	142	CONFORMANCE	10,87	3,28	3,46
56	CDP	21,18	3,98	4,82	143	MULTICAST	26,45	8,02	3,44
57	UDP	41,47	7,89	4,79	144	MAC	26,41	8,06	3,42
58	WAP	30,63	5,86	4,77	145	HDLC	7,74	2,39	3,39
59	SMTP	22,44	4,31	4,76	146	ENCRYPT	6,53	2,02	3,38
60	NCP	11,08	2,15	4,73	147	CORRECTNESS	10,87	3,38	3,37
61	LCP	17,13	3,33	4,72	148	CISCO'S	9,55	2,99	3,35
62	VINES	16,96	3,33	4,70	149	SCP	7,03	2,21	3,34
63	SPANNING	27,00	5,35	4,67	150	SLIDING	6,39	2,01	3,33
64	GATEWAY	59,48	11,80	4,67	151	MULTIHOMING	6,39	2,01	3,33
65	PPTP	26,90	5,35	4,66	152	SHIP	8,07	2,55	3,33
66	HANDSHAKE	19,30	3,84	4,66	153	TELNET	9,41	2,98	3,31
67	NLSP	14,22	2,88	4,61	154	SIGNALLING	10,39	3,35	3,26
68	SOAP	10,57	2,14	4,60	155	TTP	6,19	2,00	3,26
69	RSVP	34,54	7,04	4,59	156	ANALYZER	7,76	2,53	3,24
70	OSPF'S	13,26	2,71	4,58	157	LABEL	21,89	7,15	3,23
71	RIP	43,44	8,89	4,58	158	LAYER	51,11	16,77	3,22
72	MGCP	16,17	3,32	4,57	159	INTERNET	51,10	16,82	3,20
73	HCML	12,23	2,53	4,54	160	ZIP	7,62	2,52	3,19
74	EXTENSIBLE	20,68	4,28	4,54	161	ELECTED	6,57	2,18	3,18
75	BORDER	35,75	7,41	4,54	162	VALIDATION	12,43	4,16	3,16
76	PPP	31,94	6,69	4,51	163	AUTONOMOUS	11,52	3,87	3,15
77	RCCMP	11,13	2,34	4,50	164	EMERGED	7,70	2,65	3,08
78	TUNNELING	24,45	5,23	4,45	165	LEADER	10,53	3,63	3,07
79	SIP	25,98	5,56	4,45	166	AUTHENTICATION	22,24	7,72	3,05
80	TESTERS	9,91	2,13	4,43	167	DEFINES	16,02	5,56	3,05
81	CMIP	14,02	3,02	4,43	168	ORIENTED	17,15	5,96	3,05
82	REF	18,26	3,93	4,43	169	RELIABLE	16,57	5,76	3,05
83	DECNET	22,47	4,86	4,42	170	PROMELA	7,15	2,49	3,05
84	STACK	40,58	8,93	4,37	171	SIGNALING	16,24	5,69	3,03
85	ICMP	22,44	4,94	4,36	172	TRANSFER	23,26	8,18	3,02
86	RIB	9,53	2,12	4,33	173	UTILIZES	6,55	2,32	3,00

3.3 Definición del nodo en el diccionario.

1. Definición de *protocol* en el *Advanced Learner's Dictionary* de Cambridge (2003):

Protocol (noun):

1. *The system of rules and acceptable behaviour used at official ceremonies and occasions: a breach of Royal protocol.*
2. *A formal international agreement: The Geneva Protocol of 1925 prohibits the use of poisonous gases in war.*
3. *(Specialize) A computer language enabling computers that are connected to each other to communicate.*

2. Definición de *protocol* en el diccionario especializado Webster:

Protocol (noun): (Computer science) rules determining the format and transmission of data.

Specialty Definition:

*When used in computing, a **protocol** is a convention or standard that controls or enables the connection, communication, and data transfer between two computing devices. Protocols may be implemented by hardware, software, or a combination of the two. At the lowest level, a protocol defines a hardware connection.*

The wide-spread use and expansion of protocols is both a prerequisite to the Internet, and a major contributor to its power and success. The Transmission Control Protocol (or TCP) was the first of these, and the term TCP-IP refers to an important collection of the most used protocols. Most communication protocols are established in the RFC documents of the Internet Engineering Task Force or IETF. See also Communications protocol for more information.

Object-oriented programming has extended the use of the term to include the protocol used for connections and communication between objects.

3.4 Colocados contiguos.

a) Colocado + <i>protocol</i> .		
Total colocados seleccionados: 48	Frecuencia máxima: 173	Frecuencia mínima: 9
<ul style="list-style-type: none"> - Sustantivos, total 31: <i>internet 173, layer 127, control 112, access 111, transport 107, gateway, management 73, information 56, transfer 54, Hello 53, application 49, point 39, AppleTalk 36, datagram 34, authentication 31, resolution 31, reservation 29, micropayment 26, tree 25, distribution 24, configuration 23, message 22, communication(s) 13, initiation 17, election 16, discovery 15, tunnel 14, payment 11, level 10, security 10, delivery 9.</i> - Adjetivos, total 3: <i>new 12, specific 12, popular 9.</i> - Participios verbales que actúan como adjetivos, total 5: <i>routing 340, signaling 24/signalling 10, tunneling 23, proposed 18.</i> - Siglas, total 9: <i>OSPF 80, MAC 45, OSI 30, X 18, M 15, HTTP 12, LDAP 11, LSAP 11, ATM 9.</i> 		
b) <i>Protocol</i> + colocado.		
Total colocados seleccionados: 25	Frecuencia máxima: 28	Frecuencia mínima: 7
<ul style="list-style-type: none"> - Sustantivos, total 10: <i>unit 16/units 12, support 15, section 13, traffic 11, figure 10, background 9, structures, 8, means 8, introduction 7.</i> - Formas personales de verbo, total 4: <i>supports 9, provide 9, consists 9, defines 7.</i> - Adverbios, total 1: <i>well 7.</i> - Participios verbales que actúan como adjetivos, total 5: <i>switching 28, defined 13, transporting 7, described 7, specified 7.</i> - Participios verbales, total 1: <i>including 7.</i> - Siglas, total 4: <i>TCP 27, RFC 22, O 15, PIM 10.</i> 		

3.5 Identificación de colocados contiguos especializados en las líneas de concordancia.

*Packet Protocol (SPP). Novell also offers Internet **Protocol support** in the form of the User Datagram variable-length subnet support. The router's **IP protocol support** must include the ability to IP supernetting support. The router's **IP protocol support** must include the ability to mes unplugged. Non-broadcast lower-level **protocol support** On non-broadcast network solutions will offer dynamic, standards-based multi-**protocol support** with efficient provisioning, survival Accelerator using Cisco PXF technology Multi-**protocol support** Low initial investment Scalabil access for data centers: Complete Layer 3 Routing **Protocol Support** - Foundry's robust Layer 3 routing RIP, OSPF-TE, IS-IS(TE) and BGP-4 real-time **protocol support** The features and benefits of th of this technique are: efficient, full CMIP/SNMP **protocol support**, light and simple object model in datagrams, along with the appropriate lower-level **protocol support**, is required. The IP mul IP supernetting, see [Ref10]. Lower-level **protocol support**. The lower level protocol RIP, OSPF-TE, IS-IS(TE) and BGP-4 real-time **protocol support**. About the Routing interEMUL*

Concordancias de protocol support.

a) Colocado + protocol:

<i>access protocol</i>	<i>information protocol</i>	reservation protocol
<i>AppleTalk protocol</i>	<i>initiation protocol</i>	<i>resolution protocol</i>
<i>application protocol</i>	internet protocol	routing protocol
<i>ATM protocol</i>	<i>layer protocol</i>	<i>security protocol</i>
authentication protocol	LDAP protocol	signaling protocol
<i>communication(s) protocol</i>	<i>level protocol</i>	signalling protocol
<i>configuration protocol</i>	<i>M protocol</i>	<i>specific protocol</i>
<i>control protocol</i>	MAC protocol	transfer protocol
datagram protocol	<i>management protocol</i>	transport protocol
<i>delivery protocol</i>	<i>message protocol</i>	<i>tree protocol</i>
discovery protocol	micropayment protocol	<i>tunnel protocol</i>
<i>distribution protocol</i>	<i>multi protocol</i>	tunneling protocol
election protocol	oriented protocol	<i>X protocol</i>
gateway protocol	OSI protocol	
hello protocol	OSPF protocol	
<i>http protocol</i>	<i>point protocol</i>	

b) Protocol + colocado:

<i>protocol background</i>	<i>protocol PIM</i>	<i>protocol TCP</i>
<i>protocol figure</i>	<i>protocol RFC</i>	<i>protocol traffic</i>
<i>protocol introduction</i>	<i>protocol structures</i>	<i>protocol transporting</i>
<i>protocol means</i>	<i>protocol support(s)</i>	<i>protocol unit(s)</i>
<i>protocol O</i>	<i>protocol switching</i>	

4. AGRUPACIONES LÉXICAS O CLUSTERS.

4.1 Distribución de clusters.

Combinaciones	Volumen	Frecuencia máxima
2-cluster	526	365
3-cluster	702	32
4-cluster	439	8
5-cluster	269	26
6-cluster	146	24

4.2 Inclusión de combinaciones especializadas en unidades más amplias.

<p style="text-align: center;"><i>access protocol</i> <i>link access protocol</i> <i>access protocol LDAP</i> <i>access protocol DAP</i> <i>access protocol snap</i> <i>object access protocol</i></p>	<p style="text-align: center;"><i>communication protocol</i> <i>binary communication protocol</i> configuration protocol <i>host configuration protocol</i> <i>configuration protocol DHCP</i> <i>configuration protocol DHCP server</i></p>
---	--

<p>printer access protocol sunetwork access protocol wireless access protocol message access protocol access protocol PAP lightweight directory access protocol lightweight directory access protocol LDAP subnetwork access protocol snap header application protocol application protocol WAP wireless application protocol WAP authentication protocol extensible authentication protocol authentication protocol EAP handshake authentication protocol password authentication protocol digest authentication protocol extensible authentication protocol EAP lightweight extensible authentication protocol datagram protocol datagram protocol UDP user datagram protocol UDP delivery protocol datagram delivery protocol delivery protocol DDP datagram delivery protocol DDP discovery protocol Cisco discovery protocol discovery protocol CDP Cisco discovery protocol CDP gateway protocol gateway protocol BGP interior gateway protocol exterior gateway protocol gateway protocol EGP gateway protocol ICP gateway protocol EBGp exterior border gateway protocol multiprotocol border gateway protocol border gateway protocol EBGp IGP interior gateway protocol exterior border gateway protocol EBGp initiation protocol session initiation protocol initiation protocol SIP SIP session initiation protocol layer protocol network layer protocol transport layer protocol link layer protocol session layer protocol upper layer protocol application layer protocol layer protocol configuration layer protocol engineering layer protocol information layer protocol type track transport layer protocol layer protocol configuration negotiation network layer protocol configuration data link layer protocol network layer protocol type network layer protocol configuration negotiation standard track transport layer protocol message protocol control message protocol message protocol ICMP</p>	<p>a dynamic host configuration protocol dynamic host configuration protocol DHCP server control protocol control protocol TCP gateway control protocol link control protocol control protocol LCP session control protocol control protocol RFC control protocol SCP control protocol MGCP internet control protocol transport control protocol cache control protocol control protocol RTCP control protocol transfers special control protocol media gateway control protocol link control protocol LCP session control protocol SCP gateway control protocol MGCP TCP transmission control protocol special control protocol transfers time transport control protocol web cache control protocol media gateway control protocol MGCP distribution protocol label distribution protocol distribution protocol LDP based label distribution protocol election protocol leader election protocol information protocol information protocol ZIP zone information protocol information protocol CMIP management information protocol zone information protocol ZIP common management information protocol management information protocol CMIP RIP routing information protocol common management information protocol CMIP internet protocol control protocolinternet protocol protocolinternet protocol TCPIP IP internet protocol internet protocol VOIP internet protocol suite internet protocol version internet protocol IP networks internet protocol version (n°) voice over internet protocol VOIP voice over IP internet protocol MAC protocol control MAC protocol media access control MAC protocol medium access control MAC protocol management protocol network management protocol management protocol SNMP mobility management protocol group management protocol management protocol IGMP link management protocol key management protocol group management protocol IGMP internet group management protocol IGMP</p>
---	--

<p><i>internet control message protocol ICMP</i> micropayment protocol <i>party micropayment protocol</i> multi protocol <i>multi protocol label</i> <i>multi protocol label switching</i> <i>multi protocol label switching MPLS</i> point protocol <i>point protocol PPP</i> <i>access point protocol</i> <i>point protocol over Ethernet</i> routing protocol <i>gateway routing protocol</i> <i>state routing protocol</i> <i>OSPF routing protocol</i> <i>routing protocol packets</i> <i>routing protocol traffic</i> <i>dynamic routing protocol</i> <i>interdomain routing protocol</i> <i>QOS routing protocol</i> <i>routing protocol IGRP</i> <i>vector routing protocol</i> <i>multicast routing protocol</i> <i>routing protocol objects</i> <i>production routing protocol</i> <i>routing protocol conformance</i> <i>ATM routing protocol</i> <i>backbone routing protocol</i> <i>interconnection routing protocol</i> <i>layer routing protocol</i> <i>routing protocol EIGRP</i> <i>unicast routing protocol</i> <i>assigned routing protocol</i> <i>based routing protocol</i> <i>common routing protocol</i> <i>hierarchical routing protocol</i> <i>interior routing protocol</i> <i>routing protocol AURP</i> <i>routing protocol IDRP</i> <i>distance vector routing protocol</i> <i>gateway routing protocol EIGRP</i> <i>link state routing protocol</i> <i>network layer routing protocol</i> <i>OSPF routing protocol packets</i> <i>routing protocol conformance test</i> <i>system interconnection routing protocol</i> <i>assigned routing protocol objects</i> <i>interdomain routing protocol IDRP</i> <i>state hierarchical routing protocol</i> <i>interior gateway routing protocol EIGRP</i> <i>interior gateway routing protocol IGRP</i> <i>open system interconnection routing protocol</i> <i>primary network layer routing protocol</i> <i>routing protocol conformance test suites</i></p>	<p><i>simple network management protocol SNMP</i> OSI protocol <i>OSI protocol suite</i> OSPF protocol <i>OSPF protocol packet(s)</i> <i>OSPF protocol exchanges</i> reservation protocol <i>reservation protocol RSVP</i> <i>resolution protocol</i> <i>resolution protocol ARP</i> <i>resolution protocol AARP</i> <i>AppleTalk address resolution protocol</i> <i>AppleTalk address resolution protocol AARP</i> transfer protocol <i>file transfer protocol</i> <i>hypertext transfer protocol</i> <i>mail transfer protocol</i> <i>transfer protocol FTP</i> <i>transfer protocol HTTP</i> <i>data transfer protocol</i> <i>transfer protocol TFTP</i> <i>text transfer protocol</i> <i>transfer protocol SMTP</i> <i>file transfer protocol FTP</i> <i>trivial file transfer protocol</i> <i>HTTP hypertext transfer protocol</i> <i>SMTP simple mail transfer protocol</i> <i>trivial file transfer protocol TFTP</i> transport protocol <i>time transport protocol</i> <i>reliable transport protocol</i> <i>transport protocol class</i> <i>multicast transport protocol</i> <i>transport protocol RTP</i> <i>transport protocol header</i> <i>purpose transport protocol</i> <i>OSI transport protocol</i> <i>transport protocol VTP</i> <i>video transport protocol</i> <i>reliable multicast transport protocol</i> <i>reliable transport protocol RTP</i> <i>time transport protocol header</i> <i>transport protocol header compression</i> <i>video transport protocol VTP</i> <i>time transport protocol header compression</i> tree protocol <i>spanning tree protocol</i> tunnel protocol <i>point tunnel protocol</i> <i>tunnel protocol PTP</i> <i>layer (n°) tunnel protocol</i> <i>point tunnel protocol PTP</i></p>
--	--

4.3 Colocados contiguos de las combinaciones especializadas.

pre +	COMBINACIÓN	+ post
<p><i>directory, wireless, sub-network, object, remote, media, message, web, calendaring, link, configuration, server, printer, standard, multiple</i></p>	<p><i>access protocol</i></p>	<p><i>ARAP, -based, browser, CAP, configuration, -cordless, DAP, ELAP, FLAP, IMAP, LDAP, PAP, running, SNAP, SOAP,</i></p>
<p><i>wireless, negotiating</i></p>	<p><i>application protocol</i></p>	<p><i>anticipates, capabilities, data, incorporates, partner, approach, specification, standard, WAP</i></p>

<i>extensible, handshake, password, digest, mutual, aforementioned, actual</i>	authentication protocol	<i>CHAP, EAP, LEAP, PAP, proposed, used, verifies, version</i>
<i>binary, standardized, serial, developed, two-sided, layer, multimaster, network, main, UDP, TCP, NoC, provides, cache, popular, computer, carrier-class, proprietary, digital</i>	communication protocol	<i>designed, engineering, stack(s), WCCP, designers, developed, software, source, supporting</i>
<i>host, ES-IS</i>	configuration protocol	<i>-based, conveys, DHCP, server</i>
<i>access, VINES, cache, traffic, transmission, media-, concurrency, gateway, cache, host, time, transport, internet, terminal, link, network, IPX, bridging, DECnet, AppleTalk, IP, real-time, conference, session, efficient</i>	control protocol	<i>DCCP, defines, developed, DHCP, ICP, internet, IP, LCP, left, MGCP, processing, RTCP, SCCP, SCP, SCTP, TCP</i>
<i>user, internet</i>	datagram protocol	<i>IDP, IP, port, UDP</i>
datagram, transport	delivery protocol	<i>DDP, overview</i>
<i>Cisco, HESTIA, router-, source, service</i>	discovery protocol	<i>CDP, exposes, IDRP, IRDP, profile, SDP</i>
<i>label, tag, leader</i>	distribution protocol	<i>CR-LDP, LDP, provides</i>
	election protocol	<i>appears, discussed, informally</i>
border, interior, exterior	gateway protocol	<i>BDP, connections, eBGP, EGP, IGP</i>
routing, management, zone	information protocol	<i>CMIP, RIP, ZIP</i>
session	initiation protocol	<i>gateways, SIP</i>
<i>IP, mobile, called, over, sing, multimedia, forwarding, transmit, high-bandwidth, wireless, principally, TV, three, line, defense, offers, TCP, general-purpose, network-layer, generation, make, voice-over-</i>	internet protocol	<i>access, addresses, addressing, application, approach, -based, called, communications, data, described, -developed, device, -enabled, family, found, harmonization, instead, IP, IPv4, IPv6, journal, layer, local, multicast, networks, offering, packet(s), security, service, SLIP, stack, -structured, subnets, suite, support, version 6, truly, traffic, VoIP, VNP packets, addresses, analysis, based, configuration, defined, design, designed, developed, enables, engineering, identifiers, implementations, information, interaction(s)</i>
<i>network, higher-, presentation, link, upper-, cross- ATM, signalling, application-, transport, packet-, session,</i>	layer protocol	
<i>TEA-, stand-alone, sound, basic, actual, proposed, channel-oriented, autonomous, timeout-, sensor-, control</i>	MAC protocol	<i>allows, assigns, attempts, based, design, engine, includes, schedules, solves, used</i>
<i>mobility, distributed-, network, group, link, binding, link, data, networking, key, system, standard, layer</i>	management protocol	<i>ANMP, CGMP, code, developed, DOP, functions, IGMP, ISAKMP, LMP, SNMP, suite, used, access</i>
<i>group, communication(s), control, shared</i>	message protocol	<i>IGMP, DDCMP, ICMP, necessitate, packet, source</i>
<i>efficient, multiparty, multi-party, flipping,</i>	micropayment protocol	<i>aimed, allows, appeared, design(ed)</i>
<i>enable, new, hardware-based, consider, multi-tasking, transporting, encapsulating, streams, require, generalized, deployed, necessary, technology, standards-based, enormous</i>	multi protocol	<i>label, lambda, systems, capabilities, datagrams, encapsulation, operation, support, IGMP</i>
<i>only, entire, 7-layer, certain, all, verify, first</i>	OSI protocol	<i>class, layers, stack(s), suite(s), specifications, addresses, defines, exchanges, instance(s), makes, mechanisms, operation, packet(s), presents, requires, runs, link, used</i>
<i>base, multiple, certain, verify, first</i>	OSPF protocol	<i>addresses, defines, exchanges, instance(s), makes, mechanisms, operation, packet(s), requires, runs, link,</i>

<i>point-to-, access</i>	<i>point protocol</i>	<i>IAPP, MLP, PPP</i>
<i>resource, gateway, RSVP, time-slot</i>	<i>reservation protocol</i>	<i>BGRP, RSVP, RSVP-TE, supports, tunnelling</i>
<i>Address, hop</i>	<i>resolution protocol</i>	<i>AARP, ARP, ATM, request</i>
<i>OSPF, QoS, vector, common, underlying, backbone, system, specific, dynamic, production, scalable, update-based, gateway, interconnection, deployed, latest, include, automated, state, IPv4, DECnet, multicast, particular, link-state, hierarchical, deflection, backbone, HCML, query, distinct, interdomain, primary, dynamic, production, layer, IP, unicast, assigned, required, ATM, intra-area, internet, flow, production, hybrid, intra-domain, vector, single, zone, IPX</i>	<i>routing protocol</i>	<i>achieves, actually, addresses, algorithms, AURP, basics, chosen, components, conformance, convergence, data, developed, documented, DRP, DVMRP, EIGRP, employed, exchanges, finds, enhanced, floods, frames, IDRP, IGRP, implementation, -independent, information, interactions, involves, objects, ODMRP, packet(s), RIP, SMRP, support, traffic, updates, used, using</i>
<i>file, mail, data, hypertext</i>	<i>transfer protocol</i>	<i>FTP, HTTP, proxy, secure, SMTP, TFTP</i>
<i>multicast, data, purpose, reliable, netWare, single, optical, own, OSI, supports, real-time, new, file, simple, appropriate, realtime, flexible, wireless, rapid, TCPfriendly, mail, various, general-purpose, data, innovative, connection-oriented, multimedia, connectionless, end-to-end, news, video, upcoming</i>	<i>transport protocol</i>	<i>assures, called, class, driver, header, module, MRTP, offering, operating, provides, RCCMP, RTP, SCTP, run, sigint, stacks, traffic, UDP, used, usenet, VTP, transparency</i>
<i>spanning-</i>	<i>tree protocol</i>	<i>avoids, defines, includes, senses, STP, summary, works</i>
<i>GPRS, layer 2, point-to-point, high-performance, proprietary, IP</i>	<i>tunnel protocol</i>	<i>GTP, L2TP, operation, overview, PPTP, protocol</i>

2. Familia de *wire*. Forma analizada: WIRELESS

1. FRECUENCIA.

KEYWORD	Frec. Telec	Frec. Lacell	Ratio	Término, Chung	KEYNESS	P valor
11. WIRELESS	4.083	171	90,6825603	ESPECIFICO	11.454,00	0
Miembros relacionados:	F Telec	F Gene	Ratio	Término	Keyness	P valor
WIRED	342	110	11,8079342	NO	622,4	0
WIRE	637	611	3,95948522	NO	554	0
WIRES	288	183	5,97698137	NO	359,5	0
WIRING	209	123	6,45329738	NO	275,3	0
WIRELINE	53	0	inf/esp	inf/esp	166,2	0
WIRELESSLY	31	7	16,8191544	NO	64,2	0
LIMEWIRE	9	0	inf/esp	inf/esp	28,2	0
NANOWIRES	8	0	inf/esp	inf/esp	25,1	0,000001
NANOWIRE	4	0	inf/esp	inf/esp	12,5	0,000397
HOTWIRE	4	0	inf/esp	inf/esp	12,5	0,000397

Familia técnica: Sí, más de la mitad de los miembros son valorados como específicos.

2. DISTRIBUCIÓN.

Distribución por áreas	Distribución Keyness	Keykeyword	Nº de textos	Porcentaje
8 (-7)	4, 801, 802	WIRELESS	224	13,54%

Dispersion plot

Area	Words	Hits	Per 1,000
081 Esp. Sign	867.208	1.342	1.55
082 Esp. Tele	997.727	1.064	1.07
4 Signal proc.	580.936	603	1.04
3 Telematics	1.205.064	831	0.69
6 Business	373.079	109	0.29
1 Electronics	722.823	108	0.15
2 Ar. Comp	329.643	16	0.05
5 Materials	101.241	3	0.03

3. COLOCACIÓN.

3.1 Colocados.

N	Collocates	TOTAL	LEFT	RIGHT	L5	L4	L3	L2	L1	*	R1	R2	R3	R4	R5
1	THE	1985	1192	793	223	254	194	144	377	0	0	158	234	200	201
2	AND	1245	643	602	116	132	121	113	161	0	66	209	119	101	107
3	OF	1032	798	234	135	103	110	174	276	0	0	22	30	75	107
4	A	979	608	371	89	94	87	52	286	0	0	41	124	110	96
5	TO	963	518	445	138	107	91	138	44	0	9	132	103	99	102
6	IN	752	465	287	73	84	62	125	121	0	9	84	71	69	54
7	FOR	690	503	187	56	76	69	106	196	0	6	48	41	38	54
8	IS	559	187	372	53	62	41	24	7	0	23	124	97	68	60
9	NETWORK	409	115	294	20	33	35	25	2	0	186	25	31	22	30
10	NETWORKS	384	75	309	20	11	13	18	13	0	157	65	54	20	13
11	ACCESS	322	96	226	24	25	28	14	5	0	133	43	9	23	18

12	LAN	319	20	299	5	2	4	8	1	0	266	11	4	11	7
13	THAT	303	141	162	27	31	25	30	28	0	0	53	39	36	34
14	AS	302	163	139	46	32	32	31	22	0	0	34	43	29	33
15	COMMUNICATIONS	297	45	252	12	9	11	7	6	0	173	35	17	9	18
16	ARE	254	96	158	30	34	19	5	8	0	8	62	36	26	26
17	WITH	252	142	110	15	23	19	48	37	0	1	32	25	23	29
18	SYSTEMS	229	54	175	10	9	11	12	12	0	73	65	16	11	10
19	ON	217	135	82	26	18	20	40	31	0	0	12	19	21	30
20	TECHNOLOGY	202	50	152	18	8	10	9	5	0	88	30	13	9	12
21	MOBILE	201	99	102	14	17	14	19	35	0	32	19	17	18	16
22	DATA	200	67	133	9	17	26	12	3	0	77	12	16	11	17
23	BY	198	103	95	23	17	26	23	14	0	0	26	20	26	23
24	WILL	187	61	126	21	16	12	12	0	0	3	36	19	31	37
25	SERVICES	181	62	119	14	14	14	16	4	0	46	39	9	8	17
26	OR	178	102	76	23	18	17	27	17	0	12	24	13	14	13
27	BE	165	78	87	13	43	17	2	3	0	0	5	31	31	20
28	EQUIPMENT	165	26	139	7	6	7	4	2	0	98	20	8	8	5
29	DEVICES	155	35	120	10	8	6	8	3	0	84	16	9	8	3
30	THIS	155	54	101	10	16	9	12	7	0	0	25	32	24	20
31	CAN	150	53	97	21	10	11	9	2	0	3	35	27	18	14
32	COMMUNICATION	145	24	121	3	8	10	2	1	0	95	14	4	4	4
33	AN	143	70	73	20	25	13	12	0	0	0	9	17	28	19
34	APPLICATIONS	141	57	84	12	13	15	14	3	0	35	15	11	12	11
35	SECURITY	138	56	82	22	14	15	5	0	0	17	27	9	12	17

3.2 Colocados significativos.

Nº	Colocados	Z	T	MI	Nº	Colocados	Z	T	MI
1	YHOO	28,31	2,43	7,08	53	SBC	10,36	2,33	4,31
2	VERIO	28,31	2,43	7,08	54	INVESTING	9,33	2,12	4,28
3	CWP	28,31	2,43	7,08	55	RIM	11,47	2,67	4,20
4	BREW	26,17	2,43	6,86	56	APS	12,10	2,84	4,19
5	KERTON	24,45	2,43	6,67	57	AUCTIONS	13,74	3,27	4,14
6	CINGULAR	30,79	3,13	6,60	58	WANS	13,10	3,13	4,13
7	TFRC	61,94	6,70	6,42	59	LICENSE	20,81	5,07	4,08
8	WML	20,78	2,42	6,21	60	SPRINT	9,27	2,30	4,02
9	MULTFRC	38,90	4,73	6,08	61	WLANS	15,10	3,75	4,02
10	BELLSOUTH	25,01	3,11	6,01	62	SMC	15,82	3,97	3,99
11	SMARTPARTNER	17,27	2,20	5,94	63	PDAS	12,09	3,10	3,93
12	WECA	18,35	2,41	5,86	64	ELITECONNECT	15,79	4,07	3,91
13	BACKHAUL	20,99	2,78	5,83	65	MULTIHOP	8,00	2,08	3,88
14	WEBLINKS	16,51	2,20	5,82	66	WEP	15,80	4,16	3,85
15	TELEGRAPHY	16,51	2,20	5,82	67	ADVENT	10,57	2,79	3,84
16	SMCANT	15,83	2,19	5,70	68	BRANDS	9,17	2,46	3,80
17	VOICESTREAM	17,12	2,40	5,67	69	WI	12,13	3,33	3,73
18	ROAMABOUT	27,89	4,15	5,50	70	ADAPTERS	9,51	2,61	3,72
19	MSN	15,65	2,39	5,42	71	NETWORKING	32,43	9,08	3,67
20	OUTDOORS	14,21	2,18	5,40	72	BROADBAND	32,46	9,12	3,66
21	HOMERF	27,65	4,26	5,40	73	AP	11,17	3,18	3,62
22	TRUEPOSITION	17,83	2,76	5,38	74	BARRIER	14,59	4,20	3,59
23	WAP	42,45	6,69	5,33	75	DSSS	8,41	2,43	3,59
24	DBI	13,76	2,18	5,32	76	CORDLESS	8,89	2,59	3,56
25	LANS	69,04	10,94	5,31	77	SEAMLESS	12,06	3,54	3,54
26	HIPERLAN	17,15	2,76	5,28	78	REVOLUTION	9,34	2,74	3,54
27	SIERRA	15,87	2,58	5,25	79	WESTERN	8,22	2,42	3,53

28	MARKUP	15,24	2,57	5,13	80	EQUIPMENT	39,92	11,73	3,53
29	WIRED	50,66	8,83	5,04	81	UBIQUITOUS	9,78	2,89	3,52
30	WPAN	14,67	2,56	5,03	82	DEPLOYING	11,94	3,53	3,51
31	WPANS	12,01	2,16	4,94	83	ARENA	6,78	2,03	3,47
32	MISCELLANEOUS	14,16	2,56	4,94	84	DSL	15,74	4,80	3,43
33	PACIFIC	21,38	3,87	4,93	85	LICENSED	6,59	2,02	3,40
34	LAN	93,21	17,25	4,87	86	TELEPHONY	17,14	5,28	3,40
35	VERIZON	18,54	3,48	4,83	87	FIXED	28,74	8,94	3,37
36	WIRELIN	16,97	3,20	4,81	88	WLAN	20,59	6,44	3,35
37	EXPLOSIVE	13,47	2,55	4,80	89	NETWORK'S	6,99	2,21	3,33
38	ROGERS	16,03	3,05	4,79	90	ALLIANCE	7,98	2,54	3,30
39	HANDHELD	26,27	5,01	4,78	91	HOC	13,48	4,31	3,29
40	VTP	27,56	5,28	4,77	92	DEPLOY	11,45	3,78	3,20
41	NYSE	12,28	2,36	4,76	93	TELEPHONES	7,07	2,35	3,18
42	INDOOR	19,15	3,73	4,72	94	BRIDGES	12,25	4,08	3,18
43	FEDEX	11,67	2,35	4,62	95	OFFICIALS	6,46	2,17	3,14
44	PAGER	14,21	2,88	4,61	96	SENSOR	13,66	4,61	3,14
45	NASDAQ	15,86	3,31	4,52	97	AD	12,92	4,42	3,10
46	UWB	10,16	2,14	4,50	98	PROXIMITY	6,79	2,33	3,08
47	BLUETOOTH	39,74	8,48	4,46	99	POPULARITY	6,29	2,16	3,08
48	WTLS	9,98	2,13	4,45	100	TECHNOLOGIES	29,53	10,17	3,08
49	BLACKBERRY	14,70	3,16	4,43	101	WORLD'S	8,05	2,78	3,06
50	DECT	13,99	3,02	4,43	102	EVOLUTION	13,03	4,56	3,03
51	YAHOO	10,50	2,33	4,35	103	ADVANCES	9,95	3,50	3,01
52	OUTDOOR	14,04	3,15	4,31					

3.3 Definición del nodo en el diccionario.

1. Definición de *Wireless* en el *Advanced Learner's Dictionary* de Cambridge (2003):

Wireless (noun): (old-fashioned) a radio.

2. Definición de *wireless* en el diccionario especializado Webster:

Wireless (adj): Having no wires: A wireless security system.

Wireless (noun):

1. *Medium for communication.*
2. *Transmission by radio waves.*
3. *An electronic receiver that detects and demodulates and amplifies transmitted signals.*
4. *A communication system based on broadcasting electromagnetic waves.*

Specialty Definition:

Wireless: A term describing a computer network where there is no physical connection (either copper cable or fibre optics) between sender and receiver, but instead they are connected by radio. Applications for wireless networks include multi-party teleconferencing, distributed work sessions, personal digital assistants, and electronic newspapers. They include the transmission of voice, video, images, and data, each traffic type with possibly differing bandwidth and quality-of-service requirements. The wireless network components of a complete source-destination path requires consideration of mobility, hand-off, and varying transmission and bandwidth conditions. The wired/wireless network combination provides a severe bandwidth mismatch, as well as vastly different error conditions. The processing capability of fixed vs. mobile terminals may be expected to differ significantly. This then leads to such issues to be addressed in this environment as admission control, capacity assignment and hand-off control in the wireless domain, flow and error control over the complete end-to-end path, dynamic bandwidth control to accommodate bandwidth mismatch and/or varying processing capability.

3.4 Colocados contiguos.

a) Colocado + <i>wireless</i> .		
Total colocados: seleccionados 43	Frecuencia máxima: 75	Frecuencia mínima: 3
<ul style="list-style-type: none"> - Sustantivos, total 8: <i>generation 44, Bluetooth 41, future 17, speed 13, Verizon 12, safety 10, today's 7, Cingular 7.</i> - Adjetivos, total 10: <i>mobile 35, new 23, optical 20, public 18, pacific 15, secure 15, global 13, available 10, various 9, different 8.</i> - Participios verbales que actúan como adjetivos, total 2: <i>fixed 59, existing 14.</i> - Participios verbales, total 7: <i>using 13, based 11, leading 10, including 8, emerging 7, providing 7, oriented 7.</i> - Formas personales de verbos, total 3: <i>enable 12, deploy 10, include 7.</i> - Formas abreviadas, total 2: <i>tfr 41, ghz 7.</i> 		
b) <i>Wireless</i> + colocado.		
Total colocados seleccionados: 64	Frecuencia máxima: 266	Frecuencia mínima: 9
<ul style="list-style-type: none"> - Sustantivos, total 57: <i>network 186, communications 173/communication 95, access 133, equipment 98, technology 88/technologies 63, devices 84/device 29, data 77, systems 73/system 33, link 56/links 21, internet 47, services 46, application 42/applications 35, broadband 32, sensor 24, channel 22/channels 18, home 22, users 22, Ethernet 21, phone 20/phones 12, connectivity 19, industry 19, client 18, telephony 18, carriers 17, connection 17/connections 10, world 16, location 15, medium 15, solutions 15/solution 11, telephone 15, barrier 14, bridge 14, environment 14, knowledge 14, multimedia 14, backbone 13, evolution 13, interface 12, market 12, telecommunications 12, infrastructure 11, operators 11, routers 11/router 9, transmission 11, standards 10, nodes 9.</i> - Participios verbales, total 2: <i>networking 43, handheld 9.</i> - Adjetivos, total 2: <i>local 49, personal 10.</i> - Formas abreviadas, total 3: <i>LAN 266/LANS 93, WAN 17.</i> 		

3.5 Identificación de colocados contiguos especializados en las líneas de concordancia.

*RFACE. Intel's 82586 ISO Level 2 Ethernet **Controller Wireless** communicates with the CPU via a shared memory the controller wireless can be of help. The 82586 **controller wireless** has a "time-domain-reflectometer" command that 2 functions are implemented by the Ethernet **controller wireless** in conjunction with an Ethernet driver. The layer of 4 Mbytes. The receiving structure of the **controller wireless** is similar in some respects to the transmitting memory access unit. To mate the 82586 **controller wireless** to a transceiver, the 82501 serial interface (ESI)*

Concordancias de *wireless*.

a) Colocado + *wireless*:

<i>available wireless</i>	<i>existing wireless</i>	<i>optical wireless</i>
<i>based wireless</i>	<i>fixed wireless</i>	<i>oriented wireless</i>
<i>bluetooth wireless</i>	<i>free wireless</i>	<i>public wireless</i>
<i>cingular wireless</i>	<i>future wireless</i>	<i>safety wireless</i>
<i>deploy wireless</i>	<i>generation wireless</i>	<i>secure wireless</i>
<i>different wireless</i>	<i>ghz wireless</i>	<i>speed wireless</i>
<i>digital wireless</i>	<i>global wireless</i>	<i>tfr wireless</i>
<i>emerging wireless</i>	<i>mobile wireless</i>	<i>Verizon wireless</i>
<i>enable wireless</i>	<i>new wireless</i>	

b) *Wireless* + colocado:

<i>wireless access</i>	<i>wireless equipment</i>	<i>wireless phone(s)</i>
<i>wireless application(s)</i>	<i>wireless Ethernet</i>	<i>wireless router(s)</i>
<i>wireless backbone</i>	<i>wireless home</i>	<i>wireless sensor</i>
<i>wireless barrier</i>	<i>wireless internet</i>	<i>wireless services</i>
<i>wireless bridge</i>	<i>wireless LAN(s)</i>	<i>wireless solution(s)</i>
<i>wireless broadband</i>	<i>wireless local</i>	<i>wireless standards</i>

<i>wireless carriers</i>	<i>wireless location</i>	<i>wireless system(s)</i>
<i>wireless channel(s)</i>	<i>wireless market</i>	<i>wireless technology(ies)</i>
<i>wireless client(s)</i>	<i>wireless medium</i>	<i>wireless telecommunications</i>
<i>wireless communication(s)</i>	<i>wireless multimedia</i>	<i>wireless telephone</i>
<i>wireless connection(s)</i>	<i>wireless network(s)</i>	<i>wireless telephony</i>
<i>wireless connectivity</i>	<i>wireless networking</i>	<i>wireless transmission</i>
<i>wireless data</i>	<i>wireless nodes</i>	<i>wireless users</i>
<i>wireless device(s)</i>	<i>wireless operators</i>	<i>wireless WAN</i>
<i>wireless environment</i>	<i>wireless personal</i>	<i>wireless world</i>

4. AGRUPACIONES LÉXICAS O CLUSTERS.

4.1 Distribución de clusters.

Combinaciones	Volumen	Frecuencia máxima
2-cluster	352	50
3-cluster	647	48
4-cluster	250	6
5-cluster	60	4

4.2 Inclusión de combinaciones especializadas en unidades más amplias.

<p>bluetooth wireless <i>bluetooth wireless technology</i> <i>bluetooth wireless communication(s)</i></p> <p>digital wireless <i>digital wireless market</i> <i>digital wireless communication(s)</i> <i>digital wireless technology</i> <i>total digital wireless</i> <i>digital wireless telephone</i> <i>digital wireless communication systems</i> <i>digital wireless market today</i> <i>total digital wireless market</i> <i>digital wireless telephone systems</i></p> <p>generation wireless <i>next generation wireless</i> <i>generation wireless systems</i> <i>generation wireless LAN</i> <i>generation wireless network(s)</i> <i>third generation wireless</i> <i>fourth generation wireless</i> <i>generation wireless applications</i> <i>generation wireless LAN security</i> <i>generation wireless network operators</i> <i>third generation wireless network</i> <i>third generation wireless network operators</i></p> <p>safety wireless <i>public safety wireless</i> <i>safety wireless network</i> <i>safety wireless communications</i> <i>public safety wireless network</i> <i>public safety wireless communications</i> <i>safety wireless network PSWN</i> <i>public safety wireless network PSWN</i> <i>safety wireless network PSWN program</i></p> <p>secure wireless <i>secure wireless networks</i></p> <p>wireless access <i>fixed wireless access</i> <i>oriented wireless access</i> <i>broadband wireless access</i> <i>public wireless access</i> <i>bistsec wireless access</i> <i>oriented wireless access systems</i></p> <p>wireless client</p>	<p>fixed wireless <i>fixed wireless broadband</i> <i>broadband fixed wireless</i> <i>fixed wireless access</i> <i>spectrum fixed wireless</i> <i>fixed wireless networks</i> <i>fixed wireless bridges</i> <i>spread spectrum fixed wireless</i></p> <p>free wireless <i>free wireless backbone</i></p> <p>mobile wireless <i>mobile wireless networks</i> <i>generation mobile wireless</i> <i>mobile wireless communications</i></p> <p>new wireless <i>new wireless border</i> <i>new wireless devices</i> <i>new wireless telephony</i></p> <p>optical wireless <i>optical wireless communication(s)</i> <i>optical wireless communication systems</i></p> <p>oriented wireless <i>oriented wireless access</i> <i>oriented wireless access systems</i></p> <p>public wireless <i>public wireless access</i> <i>public wireless LAN</i> <i>public wireless network(s)</i></p> <p>speed wireless <i>high speed wireless</i> <i>speed wireless LAN</i> <i>speed wireless LAN standard</i></p> <p>wireless application(s) <i>WAP wireless application</i> <i>generation wireless applications</i> <i>WAP wireless application protocol</i> <i>coding for wireless applications</i></p> <p>wireless backbone <i>free wireless backbone</i></p> <p>wireless bridge <i>fixed wireless bridges</i></p> <p>wireless broadband <i>fixed wireless broadband</i></p> <p>wireless communication</p>
---	---

<p>the wireless client device wireless data wireless data technology pager wireless data using wireless data wireless data communications wireless data transmissions wireless data network pager wireless data service smartpartner pager wireless data smartpartner pager wireless data sevice wireless device(s) platform for wireless devices wireless devices and applications wireless Ethernet wireless Ethernet compatibility wireless home wireless home networking wireless home networks wireless home networking technologies wireless internet wireless internet access wireless internet service wireless internet service providers wireless local wireless local area wireless local loop home wireless local home wireless local area wireless local area network WLAN wireless local area networks WLANS wireless location wireless location technologies wireless market digital wireless market total digital wireless market wireless multimedia wireless multimedia communications wireless network(s) mobile wireless networks generation wireless network wireless network policy secure wireless network wireless network operators available wireless network fixed wireless network public wireless network safety wireless network wireless network access wireless network interface wireless network needs area wireless networks existing wireless networks wireless networking wireless networking equipment wireless networking options wireless networking and corporate security wireless system(s) generation wireless systems Q wireless system wireless technology(ies) bluetooth wireless technologies digital wireless technologies cost wireless technologies wireless technologies companies types of wireless technologies</p>	<p>wireless Bluetooth communication(s) digital wireless communication(s) optical wireless communication(s) indoor wireless communications mobile wireless communications safety wireless communications digital wireless communication systems optical wireless communication systems public safety communication systems family of wireless communications platforms wireless equipment testing wireless equipment enable wireless equipment duplex wireless equipment evaluate wireless equipment wireless equipment outdoors brands of wireless equipment models of wireless equipment interfaces enable wireless equipment testing wireless equipment outdoors wireless LAN wireless LAN security wireless LAN products wireless LAN access wireless LAN technology(ies) wireless LAN standard generation wireless LAN home wireless LAN wireless LAN market wireless LAN equipment public wireless LAN speed wireless LAN wireless LAN authentication wireless LAN data wireless LAN chip wireless LAN systems wireless LAN users wireless LAN vendors wireless LAN WLAN securing wireless LAN wireless LAN protocols wireless LAN services wireless LAN transmission wireless LANS WLAN friendly home wireless LANS generation wireless LAN security wireless LAN data streams wireless LAN security scheme wireless LAN security solution speed wireless LAN standard generation wireless LAN security wireless personal wireless networking area wireless networking area network(s) WPANS wireless sensor wireless sensor network(s) wireless service(s) Q wireless services wireless service providers wireless telephone wireless telephone industry wireless telephone systems digital wireless telephone digital wireless telephone systems wireless telephony wireless telephony services new wireless telephony</p>
---	---

	<i>wireless telecommunications</i> <i>wireless telecommunications equipment</i>
--	--

4.3 Colocados contiguos de las combinaciones especializadas.

PRE +	COMBINACIÓN	+ POST
<i>contrast, coolest, non-</i>	<i>bluetooth wireless</i>	<i>communication(s), connection, device, link, personal, phone, technology(ies)</i>
<i>Via</i>	<i>Cingular wireless</i>	<i>recently</i>
<i>Bluetooth, prominent, total, global, provide, Motorola's, CDMA, second-generation, TDMA</i>	<i>digital wireless</i>	<i>access, camera, MIC, communication(s), market, services, software, technology, telephone</i>
<i>delivering, spectrum, sells, longer-distance, dubbed, based, service, offering, point-to-point, broadband, high-speed, offers</i>	<i>fixed wireless</i>	<i>access, bridges, broadband, bypasses, communications, fills, internet, local, technology, metropolitan-area, networks, offers, PCS/Cellular, requires, routers, service(s), solution(s), using, vendors</i>
<i>license-</i>	<i>free wireless</i>	<i>backbone, connectivity, equipment, frequency, internet, links, network, solutions, WAN, WANS, wide-area</i>
<i>third, next-, first-, second, fourth, 3rd</i>	<i>generation wireless</i>	<i>applications, communication, LAN, mobile, multimedia, network(s), personal, spectrum, system(s), telephone</i>
<i>defining, public, secure, voicecentric, new, generation, modems</i>	<i>mobile wireless</i>	<i>access, channel, data, communication(s), device(s), GGSN, infrastructures, link, networks, operators, peers, segment, services, subscribers, telephone</i>
<i>Ericsson's, testing, developing, finding, support</i>	<i>new wireless</i>	<i>access, applications, border, broadband, capacity, devices, gadgets, LAN, platform, services, standards, telephony</i>
<i>outdoor, diffuse, indoor, multichannel, includes, hybrid</i>	<i>optical wireless</i>	<i>applications, architectures, channel, communication(s), networks, requires, system,</i>
<i>fiber-,</i>	<i>oriented wireless</i>	<i>access, systems</i>
<i>for-free, free, building, 4G</i>	<i>public wireless</i>	<i>access, hot, LAN, LANs, network(s), services, systems</i>
<i>public</i>	<i>safety wireless</i>	<i>advisory, communication(s), network, system</i>
<i>allows, providing, better, Olympic</i>	<i>secure wireless</i>	<i>communications, connectivity, data, internet, machine, networking, networks, routers, services, solution, system</i>
<i>high, higher</i>	<i>speed wireless</i>	<i>access, connectivity, equipment, LAN, network(s), provide, transmission</i>
<i>VTP and, two</i>	<i>tfrc wireless</i>	<i>behaves, perform, rate, recognizes, started, /TCP</i>
<i>runs on, visiting</i>	<i>Verizon wireless</i>	<i>advertises, corporate, jumped, said, stores</i>
<i>bit/sec, broadband, Cisco, class, different, digital, diverse, cost-effective, emerging, expandable, fibre-, fixed, flexible, focuses, heterogeneous, control-level, LMDS, makes, mobile, multihop, implement, internet, MAC-level, new, offer, openwave, fiber-oriented, overlapping, provide, public, providing, R, R2, reliable, shared, speed, standard, real-time, ubiquitous, today's, unauthorized, universal, fixed</i>	<i>wireless access</i>	<i>platform, point(s), policy, protocol, increases, PCI, PC, networks, architecture, solution, products, technologies, WAP, system(s), scenario, designs, using, protocol-cordless, channel, layer, technique,</i>

emerging, global, initial, market, using, WAP , world, next-generation, spread-spectrum, important, optical, future	wireless application(s)	protocol, opportunities, simulator, providers
license-free , licensed , multiple, need, versus	wireless backbone	links, equipment
customer's, install, typical, intros, focus, fixed ,	wireless bridge(s)	requires, configurations, contains
Canopy, fixed , full, D-link, new, offering,	wireless broadband	access, designed, products, group, system(s), solutions, applications, coverage, service, router(s), access, channels, industry, channel, technology
Aironet, enable, disassociate, links	wireless client(s)	associates, adapters , users, device, network, store, lose
Bluetooth , based, changing, non-cellular, commercial, third-generation, GHz, infrared, mobile, nowadays, range, redundant, safety, segment, 3-G, several, enable, distance, development, algorithms, allocated, concerning, digital, emerging, evolving, giving, digital, global, growing, indoor , investigate, large, leading, optical, fast-paced, particular, robust, play, secure, successful	wireless communication(s)	system(s), device(s), architecture, applications, compared, occurs, uses, fits, requirements, intended, technologies , standard(s), channels, units, platforms, exist, terminals, capabilities, series, interfaces, present, links, consume, issues, industry, field, creates, presents, infrastructure, rely, possess, cellular, solutions, functions
accessing, advanced, BellSouth , faster, GPRS, GSM, mobile, offering, outdoors , pager , proven, provide, robust, secure, specialized, two-way, voicestream's, global, using	wireless data	service(s), LAN , solution, network, technology, protocol, solutions, subscriber, logging, devices, operates, collection, link(s), communications, networking , companies, transmissions, access, coverage, connections, rate, carriers
client's, IP-enabled, mobile, Bluetooth , own, emerging, handheld , intelligent, specific, allow, burgeoning, complex, connect, first, ISPs, latest, manufacture, mobile, networks, new, oldest, battery-powered, redirect, release, today's, ubiquitous , unauthorized, portable, user, 3G, regarding, successful	wireless device(s)	using, types, PC, computers, full, research, including, began, create, utilize, cropping up, located, make, operate, need, become
broadband , different, full-duplex, enable, license-free , frequency, Ghz, IP, listing, mixing, protocols, purchase, select, selecting, specific, testing, evaluate, LOS, near, spectrum, higher-speed, evaluate, evaluating	wireless equipment	typically, family, currently, brands , vendor(s), options, operates, vendors, occasionally, solutions, contains, side, provide, allow, models, environmental, features, outdoors , measures, manufactures, software
bring, DSSS , focus, linksys, point-to-multipoint, roamabout , simultaneous, used, performance	wireless Ethernet	standard, protocol, bandwidth, adapter, PC, card, bridges , compatibility, technology
demystifying, existing, introducing, operating	wireless home	network(s), gateway, networking
enable, fixed , GPRS, including, Marconi's, offering, optical providing, secure, sites, true, weblinks ', future, PCS, future, networking	wireless internet	system, browser(s), service(s), company, access, system, protocol(s), applications, providers, computing, easier, capability, mobile
dual-band, base, example, standards-based, bit/sec, cellular, certified, client's, low-cost, early, effectively, embed, effectively, enterprise, first-generation, GHz, fast-growing, hubs, include, indoor , integrated, makes, manage, maximize, native, new, operate, securing, public,	wireless LAN(s)	medium, security, hotspots, standard, WLAN , services, protocols, architecture, access, transceiver, connections, functionality, systems, radios, environment, affected, subsystem, transmission, achieve, implementation, provides, AP , vendor, application, saves, technology(ies) , sources, equipment , products,

<i>simple, systems, high-speed, support, typical, used, undergraduate, today's, Nortel, WLAN, performance, home, hopping, installed, integrate, Cisco</i>		<i>coverage, client, base, authentication, card, adapter, solutions, capability, gear, requirements, instantly, users, products, market, data, hot, handover, connection, chip, configuration, deployments</i>
<i>access, called, enterprise-class, federal, fixed, Ghz, home, including, licensed, WLAN, WLL, broadband</i>	<i>wireless local</i>	<i>area, loop</i>
<i>accuracy, deliver, drive, cutting-edge, evaluates, global, successful</i>	<i>wireless location</i>	<i>solutions, revenue, services, technology(ies), market, industry, continues</i>
<i>3G, broadband, digital, spectrum</i>	<i>wireless market</i>	<i>acceptance, today</i>
<i>shared, collision-prone</i>	<i>wireless medium</i>	<i>takes, occupied</i>
<i>address, generation</i>	<i>wireless multimedia</i>	<i>application, networks, systems, communications, services</i>
<i>3G, airport, available, area, high-capacity, Cisco's, configure, CPE, default, different, DSSS, half-duplex, editing, enables, external, infrastructure-free, future, generation, good, multi-hop, install, integrated, local, Nortel, offer, short-range, reduce, RF, safety, high-speed, intelligent, Cingular's, Ricochet, global, entire, public, resilient, heterogeneous, ad hoc, non-, real, spectrum, switched, disparate, next-generation</i>	<i>wireless network(s)</i>	<i>throughput, offerings, policy(ies), round-trip, standard, operator, module, uses, access, interface, makes, utilizes, option, communications, applications, devices, services, barrier, operating, group, setting, card(s), PC, subscribers, network, designers, topology, connections, equipments, protection, installations, security, means, PSEN, infrastructure, barrier, features, viewpoint, routers, key, node, processing, meets, clients, performance, needs, cost, SM</i>
<i>broadband, basic, cost-effective, enterprise, home, ideal, integrated, Intel's, provide, regarding, secure, standardize, turn, built-in, peer-to-peer</i>	<i>wireless networking</i>	<i>standard, options, solution(s), equipment, company, group, products, connectivity, brings, position</i>
<i>Bluetooth, establish, third-generation, WPAN</i>	<i>wireless personal</i>	<i>digital, area, communication(s)</i>
<i>applying, low-complexity, efficient</i>	<i>wireless sensor</i>	<i>network(s), nodes</i>
<i>commercial, deploy, fixed, future, GSM, planned, providing, lifestyle-altering, 3G, expects, digital, delivering, federal, IP, making, mobile, new, offer, provide, seamless, running, waveLAN, location-based, public, secure</i>	<i>wireless services</i>	<i>provider(s), offering, needs, universal, performance, revenue, similar, today,</i>
<i>QAM-based, central, generation, international, IP, optical, safety, broadband, secure, 2G, 3G, 4G, compact, CDMA-based, cellular, competing, deploy, DS/CDMA, efficient, existing, helping, microcell, multiuser, NG, public, statewide, IPv6, fiber-oriented, underwater, packet-switching, modern, current, terrestrial</i>	<i>wireless system(s)</i>	<i>architects, transmitting, NGWS, design, standardization, advantages, used, performance, including, 4G, ride, operating, promise, operate, run, destined,</i>
<i>Bluetooth, currently, digital, cutting-edge, low-cost, exploiting, growing, includes, innovative, commercial, range, successful, users, QoS, useless, fixed, low-power, implemented, latest, leading, network-based, understand, short-range</i>	<i>wireless technology(ies)</i>	<i>further, makes, offers, companies, standard, iDen, convergence, platforms, development, designed, present, play, provide, truly, convergence, exist</i>
<i>commercial, delayed, domestic, integrated, use, personal</i>	<i>wireless telecommunication(s)</i>	<i>infrastructures, companies, technology, operations, programs, applications, base, systems, equipment, bureau</i>

<i>changing, digital, duopolistic, facto, third-generation, mobile, offer, US, cellular, currently</i>	<i>wireless telephone(s)</i>	<i>markets, networks, industry, systems, standard, service, represent</i>
<i>establish, existing, 3-G, new</i>	<i>wireless telephony</i>	<i>services, applications, appear, auctions</i>

3. Familia de *chip*. Forma analizada: CHIP

1. FRECUENCIA.

KEYWORD	Frec. Telec	Frec. Lacell	Ratio	Término, Chung	KEYNESS	P valor
154. CHIP	1.229	353	13,222625	NO	2.340,20	0
Miembros relacionados:	F Telec	F Gene	Ratio	Término	Keyness	P valor
CHIPS	392	353	4,21746867	NO	363,7	0
CHIPSET	28	4	26,585115	NO	65,6	0
CHIPSETS	15	0	inf/esp	inf/esp	47	0
MULTICHIP	11	0	inf/esp	inf/esp	34,5	0
CHIPSCOPE	7	0	inf/esp	inf/esp	22	0,000003
FASTCHIP	7	0	inf/esp	inf/esp	22	0,000003
PROCHIP	7	0	inf/esp	inf/esp	22	0,000003
MICROCHIP	19	16	4,50997487	NO	18,8	0,000014
ONCHIP	4	0	inf/esp	inf/esp	12,5	0,000397
MICROCHIPS	9	6	5,69681036	NO	10,8	0,000992

Familia técnica: Sí, más de la mitad de los miembros son valorados como específicos.

2. DISTRIBUCIÓN.

Distribución por áreas	Distribución Keyness	Keykeyword	Nº de textos	Porcentaje
9	1, 2	CHIP	114	6,84%
<i>Dispersion plot</i>				
Area	Words	Hits	Per 1,000	
2 Ar. Comp	329.643	267	0.81	
1 Electronics	722.823	447	0.62	
4 Signal proc.	580.936	172	0.30	
081 Esp. Sign	867.208	209	0.24	
6 Business	373.079	25	0.07	
3 Telematics	1.205.064	63	0.06	
7 Systems	307.691	15	0.05	
5 Materials	101.241	3	0.03	
082 Esp.Tele	997.727	31	0.03	

3. COLOCACIÓN.

3.1 Colocados.

N	Collocates	TOTAL	LEFT	RIGHT	L5	L4	L3	L2	L1	*	R1	R2	R3	R4	R5
1	THE	890	567	323	80	85	53	137	212	0	2	105	83	57	76
2	A	470	321	149	38	38	48	88	109	0	1	41	37	40	30
3	ON	355	312	43	13	15	33	102	149	0	2	8	17	9	7
4	OF	347	230	117	37	45	45	88	15	0	6	24	21	30	36
5	AND	294	123	171	28	33	24	23	15	0	30	53	29	28	31
6	TO	266	123	143	30	25	37	16	15	0	30	22	23	31	37
7	IS	214	65	149	25	22	12	6	0	0	51	20	21	30	27
8	IN	192	90	102	27	17	25	18	3	0	18	24	18	23	19
9	FOR	173	89	84	12	24	23	24	6	0	20	17	20	14	13

10	SYSTEM	109	93	16	5	3	42	39	4	0	0	8	1	3	4
11	BE	99	35	64	10	18	6	1	0	0	0	15	19	12	18
12	SINGLE	97	90	7	1	4	4	6	75	0	0	1	3	2	1
13	WITH	88	43	45	9	7	14	11	2	0	14	12	7	8	4
14	AS	86	39	47	11	7	8	10	3	0	3	13	13	11	7
15	THIS	86	32	54	12	5	4	1	10	0	0	21	13	10	10
16	ARE	79	22	57	10	7	1	2	2	0	6	16	15	9	11
17	AN	77	42	35	10	4	8	20	0	0	0	8	5	13	9
18	THAT	74	27	47	9	4	10	4	0	0	19	13	4	8	3
19	DESIGN	69	25	44	9	6	10	0	0	0	18	4	7	11	4
20	BY	56	25	31	3	7	7	8	0	0	0	5	7	9	10
21	RATE	56	8	48	0	2	2	4	0	0	44	1	1	0	2
22	CAN	50	15	35	10	3	2	0	0	0	14	5	5	6	5
23	WILL	49	7	42	2	1	4	0	0	0	10	11	5	8	8
24	AT	43	27	16	5	3	6	11	2	0	2	3	2	6	3
25	OFF	41	39	2	1	1	1	3	33	0	0	0	0	2	0
26	SYSTEMS	41	29	12	2	5	9	12	1	0	2	2	2	2	4
27	MEMORY	39	12	27	0	3	2	1	6	0	12	4	4	3	4
28	OR	39	18	21	4	7	1	4	2	0	7	3	3	6	2
29	ONE	38	20	18	3	3	0	2	12	0	1	5	4	4	4
30	WE	35	12	23	8	2	2	0	0	0	0	8	6	3	6
31	FROM	34	19	15	6	1	3	8	1	0	2	4	5	1	3
32	WHICH	34	7	27	2	2	1	1	1	0	8	10	2	6	1
33	HIGH	32	18	14	7	2	4	4	1	0	0	2	3	4	5
34	POWER	32	18	14	5	4	3	1	5	0	1	4	4	3	2
35	HAS	30	7	23	2	1	2	2	0	0	8	5	1	3	6

3.2 Colocados significativos.

Nº	Colocados	Z	T	MI	Nº	Colocados	Z	T	MI
1	MICROPHOTOGRAPH	39,98	2,23	8,33	37	TRANSCEIVER	11,38	2,52	4,36
2	XC	34,81	2,63	7,45	38	PINS	10,92	2,51	4,25
3	ETHER	26,34	2,22	7,14	39	VLSI	13,35	3,13	4,18
4	PERIPHERY	24,13	2,22	6,89	40	METHODOLOGIES	9,86	2,31	4,18
5	CDM	23,51	2,22	6,81	41	ICS	11,88	2,83	4,14
6	IMAGER	26,42	2,62	6,67	42	FPGA	20,99	4,99	4,14
7	MC	33,99	3,43	6,62	43	CORP	9,71	2,31	4,14
8	QUARTZ	21,90	2,21	6,61	44	NEURON	8,81	2,11	4,13
9	MIS	27,62	2,80	6,61	45	WCDMA	8,69	2,10	4,09
10	SPIRAL	29,09	3,13	6,44	46	RESISTORS	12,12	3,10	3,93
11	SOC	43,03	4,64	6,43	47	DISSIPATION	8,59	2,28	3,83
12	INDUCTORS	27,50	3,27	6,14	48	ONTO	14,37	3,83	3,81
13	MCM	18,20	2,20	6,09	49	TECH	7,55	2,07	3,74
14	BUMP	20,67	2,60	5,98	50	FABRICATION	9,89	2,77	3,68
15	INTERCONNECTS	28,30	3,80	5,79	51	CHIPS	10,86	3,05	3,66
16	PADS	14,83	2,19	5,52	52	INTEL	10,22	2,91	3,63
17	MEMORIES	22,34	3,38	5,45	53	DESIGNS	13,92	4,00	3,60
18	FLIP	32,61	5,07	5,37	54	IMPLEMENTS	7,62	2,24	3,54
19	TC	18,34	2,92	5,30	55	CIRCUITRY	7,56	2,24	3,52
20	MANUFACTURED	19,02	3,08	5,25	56	SILICON	14,18	4,27	3,47
21	MAKERS	15,77	2,58	5,23	57	PIN	8,50	2,57	3,45
22	SRAM	13,10	2,17	5,18	58	PARTITIONING	6,58	2,02	3,40
23	CAPACITORS	25,13	4,24	5,14	59	SPREADING	7,19	2,22	3,39
24	ATTACH	15,21	2,57	5,13	60	RAM	8,11	2,55	3,34
25	NOC	15,32	2,74	4,97	61	MM	6,35	2,01	3,32

26	FABRICATED	17,48	3,34	4,78	62	PACKAGE	11,56	3,70	3,29
27	BOND	14,01	2,72	4,73	63	WAVEFORM	7,89	2,54	3,28
28	DIE	14,80	2,89	4,72	64	TESTED	8,23	2,68	3,24
29	GOLD	14,60	2,88	4,68	65	TRANSISTORS	9,02	2,96	3,22
30	PCI	11,59	2,35	4,61	66	CMOS	9,05	2,96	3,22
31	ROM	14,74	3,03	4,57	67	GATES	8,12	2,67	3,20
32	BACKPLANE	12,29	2,53	4,56	68	XILINX	7,85	2,66	3,13
33	INTERCONNECT	17,86	3,71	4,54	69	MODULE	10,47	3,54	3,13
34	PENTIUM	10,14	2,14	4,49	70	STARTS	6,63	2,32	3,03
35	MICROPROCESSOR	11,81	2,52	4,45	71	MOTION	6,07	2,14	3,00
36	IC	17,04	3,69	4,41					

3.3 Definición del nodo en el diccionario.

- Definición de *Chip* en el *Advanced Learner's Dictionary* de Cambridge (2003):

Chip (noun):

- A long thin piece of potato that is fried and usually eaten hot: fish and chips.
- A small piece that has been broken off a larger object, or the mark left on an object such as a cup, plate, etc. where a small piece has been broken off it: wood chips. Polly fell and knocked a chip out of her front tooth.

- Definición de *chip* en el diccionario especializado Webster:

Chip (noun):

- A small fragment of something broken off from the whole; "a bit of rock caught him in the eye".
- (nautical) a triangular wooden float attached to the end of a log line.
- A piece of dried bovine dung.
- A thin crisp slice of potato fried in deep fat.
- A mark left after a small piece has been chopped or broken off of something.
- A small disk-shaped counter used to represent money when gambling.
- Electronic equipment consisting of a small crystal of a silicon semiconductor fabricated to carry out a number of electronic functions in an integrated circuit.
- A low running approach shot.
- The act of chipping something.

Specialty Definition:

In computing, the term **chip** has the following meanings:

- An integrated circuit (silicon die)
- More specifically, a microprocessor.

In telecommunications:

- In satellite communications systems, the smallest element of data in an encoded signal.
- The most elemental component of a spread spectrum signal when it is decompressed in time; that is, the longest duration signal in which signal parameters are approximately constant.
- In micrographic and display systems, a relatively small and separate piece of microform that contains microimages and coded information for search, identification, and retrieval purposes.

3.4 Colocados contiguos.

a) Colocado + <i>chip</i> .		
Total colocados: 43	Frecuencia máxima: 75	Frecuencia mínima: 3
<ul style="list-style-type: none"> - Sustantivos, total 16: <i>flip</i> 20, <i>silicon</i> 11, <i>test</i> 6, <i>controller</i> 6, <i>power</i> 5, <i>interface</i> 5, <i>ether</i> 5, <i>computer</i> 5, <i>system</i> 4, <i>prototype</i> 4, <i>project</i> 4, <i>neuron</i> 4, <i>size</i> 3, <i>quartz</i> 3, <i>Ethernet</i>, <i>code</i> 3. - Adjetivos, total 12: <i>single</i> 75, <i>one</i> 11, <i>spiral</i> 10, <i>same</i> 9, <i>specific</i> 8, <i>multiple</i> 6, <i>new</i> 5, <i>complete</i> 4, <i>full</i> 4, <i>small</i> 3, <i>first</i> 3, <i>current</i> 3. - Participios verbales que actúan como adjetivos, total 4: <i>tested</i> 5, <i>processing</i> 4, <i>manufactured</i> 4, <i>fabricated</i> 4. - Siglas, total 11: <i>FPGA</i> 18, <i>IC</i> 8, <i>VLSI</i> 6, <i>LAN</i> 6, <i>XC</i> 5, <i>MIS</i> 4, <i>GPS</i> 4, <i>CPU</i> 4, <i>ROM</i> 3, <i>DSP</i> 3, <i>SDM</i> 3. 		

b) <i>Chip</i> + colocado.		
Total colocados: 40	Frecuencia máxima: 44	Frecuencia mínima: 3
<ul style="list-style-type: none"> - Sustantivos, total 36: <i>rate</i> 44, <i>design</i> 18/<i>designs</i> 13, <i>set</i> 15/<i>sets</i> 10, <i>area</i> 14, <i>memory</i> 12/<i>memories</i> 8, <i>capacitors</i> 12, <i>inductors</i> 11, <i>size</i> 9, <i>level</i> 8, <i>verification</i> 6, <i>technology</i> 6, <i>resistors</i> 6, <i>module</i> 6/<i>modules</i> 4, <i>communication</i> 6, <i>solutions</i> 5/<i>solution</i> 3, <i>packages</i> 5, <i>makers</i> 5, <i>waveform</i> 4, <i>integration</i> 4, <i>device</i> 4, <i>companies</i> 4, <i>carrier</i> 4, <i>bond</i> 4, <i>scale</i> 3, <i>methodologies</i> 3, <i>market</i> 3, <i>manufactures</i> 3, <i>layout</i> 3, <i>duration</i> 3, <i>designers</i> 3, <i>assembly</i> 3. - Participios verbales, total 1: <i>based</i> 4. - Formas personales de verbo, total 2: <i>interconnects</i> 6, <i>implements</i> 3. - Siglas, total 1: <i>SOC</i> 16. 		

3.5 Identificación de colocados contiguos especializados en las líneas de concordancia.

RFACE. *Intel's 82586 ISO Level 2 Ethernet **Controller Chip** communicates with the CPU via a shared memory the controller chip can be of help. The 82586 **controller chip** has a "time-domain-reflectometer" command that 2 functions are implemented by the Ethernet **controller chip** in conjunction with an Ethernet driver. The layer of 4 Mbytes. The receiving structure of the **controller chip** is similar in some respects to the transmitting memory access unit. To mate the 82586 **controller chip** to a transceiver, the 82501 serial interface (ESI)*

Concordancias de *controller chip*.

a) Colocado + *chip*:

<i>code chip</i>	IC <i>chip</i>	quartz <i>chip</i>
<i>computer chip</i>	<i>interface chip</i>	ROM <i>chip</i>
<i>controller chip</i>	<i>LAN chip</i>	SDM <i>chip</i>
<i>CPU chip</i>	manufactured <i>chip</i>	silicon <i>chip</i>
<i>current chip</i>	MIS <i>chip</i>	<i>single chip</i>
<i>DSP chip</i>	<i>multi chip</i>	<i>size chip</i>
ether <i>chip</i>	<i>multiple chip</i>	<i>specific chip</i>
<i>Ethernet chip</i>	neuron <i>chip</i>	spiral <i>chip</i>
fabricated <i>chip</i>	<i>power chip</i>	<i>system chip</i>
flip <i>chip</i>	<i>processing chip</i>	<i>test chip</i>
FPGA <i>chip</i>	<i>project chip</i>	VLSI <i>chip</i>
<i>GPS chip</i>	<i>prototype chip</i>	XC <i>chip</i>

b) *Chip* + colocado:

<i>chip rate</i>	<i>chip verification</i>	<i>chip carrier</i>
<i>chip design(s)</i>	<i>chip technology</i>	<i>chip bond</i>
<i>chip SoC</i>	<i>chip resistors</i>	<i>chip solution</i>
<i>chip set</i>	<i>chip module(s)</i>	<i>chip scale</i>
<i>chip area</i>	<i>chip communication</i>	<i>chip methodologies</i>
<i>chip memory(ies)</i>	<i>chip solutions</i>	<i>chip market</i>
<i>chip capacitors</i>	<i>chip package(s)</i>	<i>chip manufacturers</i>
<i>chip inductors</i>	<i>chip makers</i>	<i>chip layout</i>
<i>chip sets</i>	<i>chip waveform</i>	<i>chip duration</i>
<i>chip size</i>	<i>chip integration</i>	<i>chip designers</i>
<i>chip level</i>	<i>chip device</i>	<i>chip assembly</i>

4. AGRUPACIONES LÉXICAS O CLUSTERS.

4.1 Distribución de clusters.

Combinaciones	Volumen	Frecuencia máxima
2-cluster	157	220
3-cluster	144	42
4-cluster	63	15
5-cluster	21	4
6-cluster	17	4

4.2 Inclusión de combinaciones especializadas en unidades más amplias.

<p>FPGA chip target FPGA chip</p> <p>MIS chip MIS chip capacitors</p> <p>multi chip multi chip module(s) multi chip module applications high reliability hybrid and multi chip hybrid and multi chip module applications high reliability hybrid and multi chip module</p> <p>quartz chip quartz chip capacitors</p> <p>specific chip specific chip device</p> <p>chip rate multiple chip rate</p> <p>chip assembly flip chip assembly</p>	<p>LAN chip wireless LAN chip LAN chip sets wireless LAN chip sets</p> <p>multiple chip multiple chip rate</p> <p>power chip power chip resistors flangeless power chip flangeless power chip resistors</p> <p>silicon chip single silicon chip</p> <p>spiral chip spiral chip inductors</p> <p>chip device specific chip device flip chip device(s) reflowing chip devices</p>
---	---

4.3 Colocados contiguos de las combinaciones especializadas.

PRE +	COMBINACIÓN	+ POST
most, probably, reflowing, five, gives, bump	flip chip	assembly(ies), bump , connection, devices, mounting, packaging, placement, resistors , technique, technology
series, types of, target wireless	FPGA chip LAN chip	using sets, include
digital, offers, carriers, DRAM , newer, SDRAM	multi chip	module(s) , solution(s), communication, devices, package(s)
run, asynchronous, DS/CDMA, neural	multiple chip	companies, generations, rate, solutions
flangeless, less	power chip	resistors
layer	quartz chip	capacitors
single, microelectronic	silicon chip	fabrication
supports	single chip specific chip	being, demodulator, related, implementation, -in- package , device
USM, Q, all, custom, microwaves	spiral chip	inductor(s)
uses, CDM , code, fixed, high, lowest, multiple, nominal, variable, higher, common, interface	chip rate	reduced, sampling, code, MCR, DS/CDMA, RBF, during, assuming, VCR, symbol, lower, optical, reconfiguration,
flip	chip assembly(ies)	line
specific, flip , multi-, system-on-system-on-a-, system-on-, pacific, single-, standardize, system-, today's, complete, compatible, custom, demodulator, processor, test, made, signal, cellular, current	chip device(s) chip design(s)	requires, MCDs, using today, language(s), methodologies , process, starts, sit up, worldwide, working, enabled
off-, on-, system-on-, own	chip memory(ies)	units, controller, whilst, distributed, circuits
CDM , code, fixed, high, lowest, multiple, original, nominal, variable, higher, common, interface	chip rate	structure, sampling, code, MCR, RBF, VCR, symbol, lower, optical
SPARC-2, CPU, dedicated, front-end, Ethernet, CN400, decoder, support, lance, complete, covers, JPEG, ATM, WLAN, LAN	chip set(s)	architecture, systems, comes, size, meets, begins
full-, system-on-a-	chip SoC	design(s) , prototypes, ICs , test, integrated, options, implementation, devices,

4. Familia de *multicast*. Forma analizada: MULTICAST

1. FRECUENCIA.

KEYWORD	Frec. Telec	Frec. Lacell	Ratio	Término, Chung	KEYNESS	P valor
133. MULTICAST	837	0	inf/esp	inf/esp	2.625,30	0
Miembros relacionados:	F Telec	F Gene	Ratio	Término	Keyness	P valor
MULTICASTING	70	0	inf/esp	inf/esp	219,5	0
MULTICASTS	29	0	inf/esp	inf/esp	91	0
Familia técnica: Sí, todos los miembros son valorados como específicos.						

2. DISTRIBUCIÓN.

Distribución por áreas	Distribución Keyness	Keykeyword	Nº de textos	Porcentaje
8	802	MULTICAST	41	2,48%
<i>Dispersion plot</i>				
Area	Words	Hits	Per 1,000	
082 Esp. Tele	997.727	527	0.53	
3 Telematics	1.205.064	174	0.14	
081 Esp. Sign	867.208	97	0.11	
7 Systems	307.691	12	0.04	
2 Ar. Comp	329.643	7	0.02	
1 Electronics	722.823	15	0.02	
6 Business	373.079	3	0.01	
4 Signal proc.	580.936	2	0.00	

3. COLOCACIÓN.

3.1 Colocados.

N	Collocates	TOTAL	LEFT	RIGHT	L5	L4	L3	L2	L1	*	R1	R2	R3	R4	R5
1	THE	483	277	206	46	61	34	46	90	0	1	46	65	44	50
2	A	299	191	108	19	32	24	34	82	0	0	14	40	31	23
3	TO	268	166	102	20	37	54	51	4	0	9	20	14	26	33
4	AND	184	105	79	27	11	15	13	39	0	7	38	11	8	15
5	IS	169	71	98	22	25	14	7	3	0	8	27	26	19	18
6	OF	166	137	29	19	11	26	43	38	0	0	6	6	5	12
7	IN	131	52	79	8	15	11	14	4	0	6	32	12	19	10
8	FOR	129	89	40	6	9	16	38	20	0	2	9	6	13	10
9	IP	98	86	12	2	6	5	0	73	0	3	3	2	1	3
10	ROUTING	97	25	72	9	1	7	6	2	0	57	2	3	7	3
11	ARE	92	41	51	7	9	11	5	9	0	0	22	12	8	9
12	GROUP	88	13	75	0	4	4	1	4	0	62	3	2	5	3
13	TREE	86	21	65	11	7	2	1	0	0	43	13	1	0	8
14	THAT	85	42	43	17	5	6	9	5	0	0	19	6	10	8
15	ADDRESS	84	23	61	8	5	6	4	0	0	42	12	1	2	4
16	PROTOCOL	79	38	41	7	7	2	18	4	0	4	22	4	6	5
17	PACKETS	70	33	37	11	9	5	6	2	0	19	5	7	2	4
18	ON	64	26	38	4	8	8	6	0	0	3	9	9	10	7
19	BE	60	34	26	12	11	2	1	8	0	0	1	13	6	6
20	NETWORK	58	28	30	6	3	5	12	2	0	4	0	8	9	9

21	ADDRESSES	56	7	49	3	0	3	0	1	0	31	6	3	7	2
22	OR	56	29	27	5	4	3	1	16	0	10	9	3	3	2
23	AS	55	31	24	7	5	11	3	5	0	0	3	7	6	8
24	THIS	55	36	19	9	4	10	2	11	0	0	6	2	5	6
25	WITH	55	33	22	6	10	7	3	7	0	2	7	4	3	6
26	TRAFFIC	53	9	44	1	6	0	2	0	0	36	3	3	1	1
27	UNICAST	51	29	22	2	1	4	15	7	0	0	9	6	5	2
28	BROADCAST	49	26	23	2	1	1	16	6	0	2	9	7	3	2
29	ALL	47	16	31	2	5	3	2	4	0	1	5	7	10	8
30	RELIABLE	46	44	2	2	0	1	1	40	0	0	0	0	2	0
31	BY	41	22	19	5	2	5	7	3	0	2	2	6	5	4
32	IT	37	11	26	6	1	4	0	0	0	0	14	1	5	6
33	DATA	36	17	19	4	2	8	3	0	0	7	1	7	0	4
34	FORWARDING	32	17	14	5	2	1	1	8	0	7	2	0	3	2
35	ROUTER	30	15	15	6	2	7	0	0	0	4	0	5	3	3

3.2 Colocados significativos.

Nº	Colocados	Z	T	MI	Nº	Colocados	Z	T	MI
1	ALLSPFRouters	69,64	2,99	9,08	26	CONSTRUCTING	9,54	2,13	4,33
2	ALLDRouters	45,34	2,23	8,69	27	FEC	10,79	2,50	4,22
3	MOSPF	52,92	2,99	8,29	28	LEAVE	9,93	2,32	4,20
4	PIM	86,94	5,36	8,04	29	DATAGRAM	9,03	2,11	4,19
5	SMRP	73,35	4,67	7,95	30	RSVP	10,49	2,50	4,14
6	UNICAST	102,22	7,11	7,69	31	RFC	11,66	2,82	4,09
7	PRAGMATIC	41,35	2,98	7,58	32	RECEIVERS	12,84	3,24	3,97
8	ANYCAST	30,14	2,22	7,52	33	MAPPED	7,70	2,07	3,78
9	RCCMP	22,87	2,22	6,74	34	HELLO	10,88	2,93	3,78
10	DATAGRAMS	42,55	4,20	6,68	35	TAG	7,61	2,24	3,53
11	JOINS	23,27	2,42	6,53	36	UPDATES	8,75	2,58	3,52
12	IGMP	25,20	2,79	6,35	37	ADDRESSED	7,90	2,40	3,43
13	TREES	43,90	4,94	6,31	38	EXTENSIONS	6,50	2,02	3,37
14	MULTICASTING	24,26	2,79	6,24	39	HOSTS	8,69	2,71	3,36
15	VCC	29,43	3,42	6,21	40	NORMALLY	6,96	2,20	3,32
16	CLASSIFY	16,59	2,20	5,83	41	TRANSACTION	6,34	2,01	3,31
17	RELIABLE	47,05	6,65	5,65	42	MEMBERSHIP	6,90	2,20	3,30
18	LISTEN	16,83	2,58	5,41	43	SHAPE	6,87	2,20	3,29
19	BROADCAST	42,45	6,82	5,28	44	SCOPE	8,41	2,69	3,29
20	SESSIONS	24,06	4,01	5,17	45	ADDRESSING	7,34	2,37	3,26
21	FORWARDING	29,88	5,39	4,94	46	WDM	7,32	2,37	3,25
22	JOIN	16,03	2,90	4,93	47	UNDERLYING	7,79	2,53	3,24
23	KEEPS	12,48	2,36	4,80	48	BUILD	7,97	2,78	3,04
24	FORWARDED	13,63	2,72	4,65	49	SENDING	7,51	2,63	3,02
25	SPANNING	10,33	2,14	4,54					

3.3 Definición del nodo en el diccionario.

1. El *Advanced Learner's Dictionary* de Cambridge (2003) no registra la definición de *multicast*. El lema más cercano a *multicast* en este diccionario es *cast*:

Cast (verb):

1. (light) to send light or shadow (= an area of darkness) in a particular direction.
2. (throw) to throw something.
3. (actors) to choose actors to play particular parts in a play, film or show.
4. (shape) to make an object by pouring hot liquid, such as melted metal, into a shaped container where it becomes hard.

5. (vote) to vote.
6. (skin) If a snake casts its skin, the outer layer of old skin comes off its body.

2. Definición de *multicast* en el diccionario especializado Webster:

Multicast Specialty Definition: *Multicast is a delivery of information to multiple destinations simultaneously. Typically used to refer to IP Multicasting, which is a protocol for efficiently sending to multiple receivers at the same time on TCP-IP networks, by use of a multicast address. It's also commonly associated with audio/video protocols such as RTP.*

3.4 Colocados contiguos.

a) Colocado + <i>multicast</i> .		
Total colocados seleccionados: 14	Frecuencia máxima: 39	Frecuencia mínima: 4
<ul style="list-style-type: none"> - Sustantivos, total 2: <i>layer 13, support 5</i>. - Adjetivos, total 7: <i>reliable 39, independent 15, particular 9, different 7, general 4, local 4, special 4</i>. - Participios verbales, total 2: <i>forwarding 8, underlying 4</i>. - Formas personales de verbos, total 1: <i>receive 5</i>. - Siglas, total 2: <i>IP 68, SMRP 9</i>. 		
b) <i>Multicast</i> + colocado.		
Total colocados seleccionados: 19	Frecuencia máxima: 22	Frecuencia mínima: 3
<ul style="list-style-type: none"> - Sustantivos, total 17: <i>protocol 22, tree 13/tress 4, address 12, sessions 6, transport 6, protocols 5, satellite 5, allspfrouters 4, management 4, mode 4, packet 4, use 4, anycast 3, class 3, delivery 3, problem 3</i>. - Participios verbales, total 1: <i>sending 3</i>. - Siglas, total 1: <i>VVC 7</i>. 		

3.5 Identificación de colocados contiguos especializados en las líneas de concordancia.

...y take large cost. The Fig. 1 shows an original **multicast tree**, and Fig. 2 draws a possible con-
MPLS domain. • Maintaining the shape of the **multicast tree**, so no modification on the multic-
ch as an RSVP flow), or it could be bound to a **multicast tree**. The control component creates
ch as an RSVP flow), or it could be bound to a **multicast tree**. The control component is resp-
sed to calculate the number assigned for each **multicast tree**. The calculated number is used
ith splitting capability is useful in expanding the **multicast tree**. However, a node with only splitt-
address of each LER and LSR spanned by the **multicast tree**. The number of entries in the TN
erated by the ingress LER and the LSRs of the **multicast tree**. Reducing the memory size mea-
IP addresses of the all routers spanned by the **multicast tree**. In [11], we have proposed an a-
es are kept only in the branching routers of the **multicast tree**. Then, LSPs are built only betw-
h has more than one downstream on the WDM **multicast tree**. We show that these additional f-
bust control tree that keeps a close shape with **multicast tree**. The proposed mechanism doe-
ways exist to create binding between tags and **multicast trees** (routes). For a set of tag switch-
in [8] have proposed an approach to construct **multicast trees** in MPLS networks. In their sc-
different MPLS label. However, the number of **multicast trees** that could be built over the net-
for building QoS-assured, delay-efficient service **multicast trees** that uses geometric location inf-
PLS network with a number of LERs. Several **multicast trees** could be built for each ingress-
ee that connects the group members together. **Multicast trees** can fit into two categories (1)
se a protocol to construct source-rooted WDM **multicast trees**. The protocol works under dens-

Concordancias de *multicast tree(s)*.

a) Colocado + *multicast*:

<i>IP multicast</i>	<i>particular multicast</i>	<i>general multicast</i>
reliable <i>multicast</i>	SMRP <i>multicast</i>	<i>local multicast</i>
<i>independent multicast</i>	forwarding <i>multicast</i>	<i>special multicast</i>
<i>layer multicast</i>	<i>support multicast</i>	underlying <i>multicast</i>

b) Multicast + colocado:

<i>multicast protocol</i>	<i>multicast protocols</i>	<i>multicast anycast</i>
<i>multicast tree(s)</i>	<i>multicast satellite</i>	<i>multicast class</i>
<i>multicast address(es)</i>	<i>multicast allspfrouters</i>	<i>multicast delivery</i>
<i>multicast VCC</i>	<i>multicast management</i>	<i>multicast problem</i>
<i>multicast sessions</i>	<i>multicast mode</i>	<i>multicast sending</i>
<i>multicast transport</i>	<i>multicast packet</i>	

4. AGRUPACIONES LÉXICAS O CLUSTERS.

4.1 Distribución de clusters.

Combinaciones	Volumen	Frecuencia máxima
2-cluster	103	97
3-cluster	144	22
4-cluster	75	8
5-cluster	34	4

4.2 Inclusión de combinaciones especializadas en unidades más amplias.

<p>IP multicast <i>IP multicast address(es)</i> <i>IP multicast datagrams</i> <i>IP multicast group</i> <i>IP multicast traffic</i> independent multicast <i>independent multicast PIM</i> layer multicast <i>layer multicast addresses</i> <i>link layer multicast</i> <i>network layer multicast</i> <i>data link layer multicast</i> <i>network layer multicast addresses</i> <i>link layer multicast addresses</i> forwarding multicast <i>forwarding multicast datagrams</i> <i>forwarding multicast traffic</i> multicast address(es) <i>layer multicast addresses</i> <i>IP multicast address(es)</i> <i>SMRP multicast addresses</i> <i>network layer multicast addresses</i> <i>link layer multicast addresses</i> <i>the multicast address allspfrouters</i></p>	<p>reliable multicast <i>reliable multicast transport</i> <i>reliable multicast protocols</i> <i>reliable multicast service</i> <i>mechanism for reliable multicast</i> <i>reliable multicast transport protocol</i> <i>reliable multicast over satellite</i> <i>reliable multicast transport service</i> <i>configuration mechanism for reliable multicast</i> <i>reliable multicast over satellite networks</i> particular multicast <i>a particular multicast group</i> SMRP multicast <i>SMRP multicast addresses</i> local multicast <i>local multicast router</i> underlying multicast <i>underlying multicast routing</i> <i>underlying multicast routing tree</i> <i>match the underlying multicast routing tree</i> <i>underlying multicast routing tree topology</i> multicast transport <i>reliable multicast transport</i> <i>reliable multicast transport protocol</i> <i>reliable multicast transport service</i></p>
--	---

4.3 Colocados contiguos de las combinaciones especializadas.

PRE +	COMBINACIÓN	+ POST
<i>nodes</i>	forwarding multicast	datagrams, packets, traffic
<i>protocol-, completely</i>	independent multicast	called, PIM, routing
<i>current, particular, address, special, distinct, available, utilise, difficult for, forward, sending, using, all-systems, specific, Cisco, supports, scalable, allow, routing, protocol</i>	IP multicast	<i>provides, address(s), connections, datagrams, delivers, does, group, IP, makes, options, packet, routing, services, sessions, standards, technology, traffic, UPD</i>
<i>network, transport, data-link</i>	layer multicast	<i>address(es), capabilities, technique</i>
	local multicast	router
<i>enable, most, current, tree-based, TCP-friendly</i>	reliable multicast	<i>85, 89, over satellite, protocols, service(s), transport</i>
	SMRP multicast	address(es), group, routes, transaction, transport

	underlying multicast	routing
<i>IP, layer, manage, SMRP, special, AllEndpoints, All-SPFRouters, reserved, standard, several, layer2</i>	multicast address(es)	allocation, recognition, AllSPFRouters, assignment, mapping, dynamic, management, fulfil, -used
	multicast forwarding	VCC, states, trees, function, tables
<i>change, different, IP, particular, per-, SMRP, unique, virtual, concerned, allows, leave</i>	multicast group(s)	address(es), set up, routing, IDs, membership , labels, concept, management, participation, specified, efficiently, use, establish
<i>IP, every, single, exchange, forwarding, sending, Q-</i>	multicast packet(s)	arrives, loss, addressed , over network, quickly
<i>group, controlled, reliable, following, supporting, new, popular</i>	multicast protocol(s)	IGMP, PIM-SIM , targets, use, resulting, -RIP
<i>one-to-many, reliable, based, furnish, provide, specifically, IP</i>	multicast service(s)	
<i>classify, corresponding, sat-RMTP, ESVP, IP, different, entire</i>	multicast session(s)	sends, using, delivered, pass, increases
<i>constrained, flood(s), forward, forwarding, IP, stream of, volume of, prevent, receive, routing</i>	multicast traffic	management, via, takes, follows,
<i>appropriate, reliable, SMRP</i>	multicast transport	protocol(s), service(s), RMT
<i>common, each, given, particular, possible, original, WDM, construct, service, several</i>	multicast tree(s)	passes, is built, allocates, rooted, built, routes

5. Familia de *cache*. Forma analizada: CACHE

1. FRECUENCIA.

KEYWORD	Frec. Telec	Frec. Lacell	Ratio	Término, Chung	KEYNESS	P valor
163 CACHE	781	20	148,306963	ESPECIFICO	2.271,90	0
Miembros relacionados:	F Telec	F Gene	Ratio	Término	Keyness	P valor
CACHING	142	2	269,649024	ESPECIFICO	425,2	0
CACHES	119	4	112,986739	ESPECIFICO	339,8	0
CACHED	71	0	inf/esp	inf/esp	222,7	0
Familia técnica: Sí, todos los miembros son valorados como específicos.						

2. DISTRIBUCIÓN.

Distribución por áreas	Distribución Keyness	Keykeyword	Nº de textos	Porcentaje
9	1, 2, 082	CACHE	30	1,81%
<i>Dispersion plot</i>				
Area	Words	Hits	Per 1,000	
2 Ar. Comp	329.643	145	0.44	
1 Electronics	722.823	208	0.29	
7 Systems	307.655	74	0.24	
082 Esp. Tele	997.727	0230	0.23	
4 Signal proc.	580.936	52	0.09	
081 Esp. Sign	867.208	43	0.05	
6 Business	373.079	12	0.03	
5 Materials	101.241	2	0.02	
3 Telematics	1.205.064	16	0.01	

3. COLOCACIÓN.

3.1 Colocados.

N	Collocates	TOTAL	LEFT	RIGHT	L5	L4	L3	L2	L1	*	R1	R2	R3	R4	R5
1	THE	547	346	201	42	48	55	87	114	0	8	36	57	44	56
2	A	261	165	96	22	18	24	34	67	0	0	13	33	24	26
3	TO	219	131	88	29	24	28	35	15	0	12	12	21	23	20
4	AND	207	87	120	21	25	19	6	16	0	19	46	18	15	22
5	OF	184	118	66	18	14	15	36	35	0	5	16	10	16	19
6	IN	142	88	54	18	11	21	28	10	0	6	21	6	12	9
7	IS	112	46	66	18	13	7	7	1	0	10	18	18	13	7
8	MEMORY	95	38	57	13	8	6	8	3	0	25	5	10	12	5
9	ENGINE	94	9	85	1	3	3	2	0	0	81	0	0	2	2
10	FOR	89	44	45	12	4	8	17	3	0	7	16	7	6	9
11	DATA	83	39	44	7	8	7	5	12	0	1	27	2	6	8
12	THAT	76	35	41	5	10	5	6	9	0	2	8	12	11	8
13	CONSCIOUS	58	2	56	1	1	0	0	0	0	55	0	0	0	1
14	ARE	55	20	35	8	9	3	0	0	0	5	12	3	5	10
15	LEVEL	50	46	4	4	2	5	6	29	0	0	0	1	0	3
16	CAN	47	18	29	7	2	5	3	1	0	2	12	7	6	2
17	FROM	47	26	21	3	4	4	12	3	0	3	4	6	5	3

18	WITH	46	24	22	4	7	3	8	2	0	3	6	2	5	6
19	IF	44	19	25	4	4	4	7	0	0	1	13	5	2	4
20	CLUSTER	43	14	29	6	3	0	3	2	0	12	7	0	8	2
21	THIS	42	18	24	7	8	2	0	1	0	0	5	6	7	6
22	L	39	35	4	2	0	2	3	28	0	0	1	1	1	1
23	AS	38	15	23	4	0	2	6	3	0	1	9	3	7	3
24	BE	38	19	19	3	15	1	0	0	0	0	2	9	3	5
25	NETWORK	36	21	15	5	2	4	4	6	0	0	4	5	6	0
26	BY	35	18	17	3	4	4	6	1	0	1	7	3	2	4
27	ON	33	10	23	1	2	0	7	0	0	1	9	3	6	4
28	AN	32	18	14	2	6	3	7	0	0	0	1	6	4	3
29	BLOCK	32	1	31	0	1	0	0	0	0	27	0	1	1	2
30	ENGINES	32	3	29	2	0	0	1	0	0	28	0	0	1	0
31	IT	31	13	18	4	4	5	0	0	0	0	8	6	3	1
32	NOT	31	14	17	4	4	3	2	1	0	0	3	7	3	4
33	PERFORMANCE	31	13	18	3	2	6	2	0	0	10	0	2	3	3
34	MISSES	30	4	26	0	0	1	0	3	0	24	0	0	1	1
35	WEB	29	19	10	1	7	0	2	9	0	1	2	2	2	3

3.2 Colocados significativos.

N°	Colocados	Z	T	MI	N°	Colocados	Z	T	MI
1	SEWD	76,63	2,23	10,21	29	SIZES	26,93	4,46	5,20
2	MISSES	124,23	5,47	9,02	30	PAGING	14,40	2,38	5,20
3	CONSCIOUS	162,07	7,60	8,83	31	PIPELINE	14,33	2,38	5,18
4	PROCESSOR'S	38,22	2,23	8,21	32	RATIOS	12,10	2,16	4,97
5	KBYTES	58,00	3,45	8,15	33	ACCESSES	14,18	2,56	4,95
6	ALLOCATOR	45,14	2,82	8,01	34	CONFLICT	12,98	2,37	4,91
7	WCCP	43,20	2,82	7,88	35	SEMI	12,17	2,36	4,74
8	POLLUTION	34,61	2,44	7,66	36	MAPPED	14,85	2,89	4,73
9	BUCKETS	34,61	2,44	7,66	37	CLUSTERING	12,72	2,54	4,65
10	COHERENCE	60,62	4,45	7,54	38	TRANSPARENT	16,75	3,45	4,56
11	ENGINES	70,40	5,62	7,30	39	ENTRIES	14,45	3,29	4,28
12	ENGINE	117,20	9,63	7,22	40	PROXY	15,52	3,55	4,26
13	REDIRECTED	31,68	2,63	7,19	41	BLOCKING	9,87	2,32	4,19
14	LOCALITY	44,80	3,72	7,19	42	REQUESTS	18,01	4,23	4,19
15	MISS	51,70	4,33	7,16	43	TREES	8,65	2,10	4,08
16	HEAP	37,25	3,43	6,88	44	BALANCING	8,01	2,09	3,89
17	CLUSTERED	26,04	2,62	6,63	45	BINDING	8,29	2,27	3,75
18	EVENLY	24,06	2,42	6,63	46	HIERARCHY	9,36	2,61	3,69
19	CONFLICTS	29,37	3,13	6,47	47	ADDS	7,77	2,25	3,59
20	COLORED	22,75	2,42	6,47	48	SUN	7,73	2,24	3,57
21	CACHING	33,00	3,82	6,23	49	LATENCY	11,35	3,30	3,57
22	HIT	31,13	3,82	6,06	50	AUTOMATIC	8,80	2,59	3,54
23	BYPASS	25,18	3,11	6,04	51	SERVE	7,31	2,22	3,44
24	CACHES	21,51	2,94	5,74	52	INSTRUCTION	8,75	2,71	3,38
25	PLACEMENT	33,62	4,60	5,74	53	REMAINING	6,43	2,02	3,35
26	TRANSPARENTLY	14,71	2,19	5,51	54	ENABLED	7,30	2,37	3,25
27	ARP	21,10	3,24	5,41	55	RATIO	8,60	2,82	3,22
28	FOOTPRINT	13,89	2,18	5,35	56	CLIENTS	6,68	2,33	3,05

3.3 Definición del nodo en el diccionario.

1. Definición de *Cache* en el *Advanced Learner's Dictionary* de Cambridge (2003):

Cache (noun): a hidden store of things, or the place where they are kept: an arms cache.

Cache (memory): an area or type of computer memory in which information that is often in use can be stored temporarily and accessed especially quickly.

2. Definición de *cache* en el diccionario especializado Webster:

Cache (noun):

1. A hidden storage space (for money or provisions or weapons).
2. A secret store of valuables or money.
3. (computer science) RAM memory that is set aside as a specialized buffer storage that is continually updated; used to optimize data transfers between system elements with different characteristics.

Cache (verb): Save up as for future use.

Specialty Definition:

In computer science: A cache is a tiny short-term memory space in a computer with quicker access speed than main memory to fill the gap between slow devices and fast devices. There may be multiple levels of cache, and in general such an architecture is referred to as a hierarchical memory system.

3.4 Colocados contiguos.

a) Colocado + <i>cache</i> .		
Total colocados seleccionados: 25	Frecuencia máxima: 29	Frecuencia mínima: 3
<ul style="list-style-type: none"> - Sustantivos, total 7: <i>level 29, Cisco 15, name 14, web 9, proxy 8, instruction 6, network 6.</i> - Adjetivos, total 9: <i>new 8, different 7, local 7, same 5, single 5, two 5, automatic 4, large 4, transparent 3.</i> - Formas personales de verbos, total 1: <i>reduce 6.</i> - Participios verbales, total 6: <i>based 5, binding 5, routing 5, mapped 5, integrated 3, paging 3.</i> - Siglas, total 1: <i>ARP 10, SEWD 3.</i> 		
b) <i>Cache</i> + colocado.		
Total colocados seleccionados: 28	Frecuencia máxima: 81	Frecuencia mínima: 4
<ul style="list-style-type: none"> - Sustantivos, total 24: <i>engine 81/engines 28, block 27, memory 25, coherence 18, sizes 18/ size 10, cluster 12, blocks 10, locality 10, performance 10, conflicts 8, control 8, hit 7, entries 6, model 6, parameters 6, capacity 5, entry 5, line5/lines 5, configuration 4, footprint 4, server 4.</i> - Adjetivos, total 1: <i>conscious 55.</i> - Formas personales de verbo, total 3: <i>misses 24/miss 12, read 4.</i> 		

3.5 Identificación de colocados contiguos especializados en las líneas de concordancia.

present a technique for determining the best data cache size required for a given memory-intensive cache sizes. This does not mean that, in general, cache size is a parameter of no impact on performance response time reduction with increased memory cache size for Web servers. However, such modification only to say that for the selected benchmarks, the cache size is not effective. 6. Conclusion. In this concerning power consumption (by determining the cache size), main memory size (by varying the data effects a significant reduction in the required data cache size, with no negative impact on the performance represent design decisions, such as clock speed, cache size, and operation bit width. The interconnect is commonly caused by an increased available cache size. In previous work [8], we presented a number of parameters is exported to the user, such as the cache size. Modifications that affect the instructed to the designer in the problem space, such as cache sizes and the clock frequency. These parameters run using mesh (or ring) network for second level cache sizes of 256 and 4048 KBytes, and so would be due to either the impact of the second level cache sizes or the selected benchmarks. Note that cache sizes of 8 and 16 KBytes, and second level cache sizes of 256 and 4048 KBytes. Fig. 6, Fig. latency, simulations are carried out for first level cache sizes of 8 and 16 KBytes, and second level

with shared-storage architectures. 5.3. Impact of **cache sizes** on latency. To investigate the impact of **cache sizes** on memory latency, simulations are run and Fig. 9. Fig. 8 and Fig. 9 show the impact of **cache sizes** on overall memory latency for mesh 8-cluster system the performance depends on the **cache sizes** as shown in Fig. 7. When cache size increases the overall memory latency decreased when the **cache sizes** were increased from 8 to 16 KBytes for mesh 8-cluster system the performance depends on the **cache sizes** as shown in Fig. 7. When **cache sizes** increase the memory latency decreases with the increase of first level and second level **cache sizes**. Note that CL1 and CL2 refer to first and second level cache sizes. For some specific cases, this change in memory latency decreases with the increase of **cache sizes**. Fig. 8. Memory latency (in 1000 processor clocks) of 8 cluster ring topology with variable **cache sizes**. Fig. 9. Memory latency (in 1000 processor clocks) of 8 cluster hypercube topology with variable **cache sizes**. Mesh, hypercube, and ring network topologies of 8 cluster mesh topology with variable **cache sizes**. Fig. 9. Memory latency (in 1000 processor clocks), which remained almost constant for various **cache sizes**. This does not mean that, in general,

Concordancias de cache size(s).

a) Colocado + cache:

ARP cache	level cache	SEWD cache
automatic cache	local cache	routing cache
binding cache	mapped cache	single cache
Cisco cache	name cache	transparent cache
different cache	network cache	web cache
instruction cache	paging cache	
integrated cache	proxy cache	

b) Cache + colocado:

cache block(s)	cache control	cache memory
cache capacity	cache engine (s)	cache model
cache cluster	cache entry (ies)	cache parameters
cache coherence	cache footprint	cache performance
cache configuration	cache hit	cache read
cache conflicts	cache line(s)	cache server
cache conscious	cache locality	cache size (s)

4. AGRUPACIONES LÉXICAS O CLUSTERS.

4.1 Distribución de clusters.

Combinaciones	Volumen	Frecuencia máxima
2-cluster	86	20
3-cluster	113	24
4-cluster	57	17
5-cluster	18	5

4.2 Inclusión de combinaciones especializadas en unidades más amplias.

<p>automatic cache automatic cache- conscious automatic cache- conscious data automatic cache- conscious data placement Cisco cache Cisco cache engine(s) integrated cache integrated cache engine level cache first level cache second level cache level cache sizes level cache read first level cache read second level cache sizes cache control cache control protocol</p>	<p>mapped cache direct mapped cache name cache name cache entries netbios name cache single cache single cache clock web cache web cache control web cache control protocol cache block(s) single cache block cache block size clustered in a cache block cache engine(s) Cisco cache engines integrated cache engine</p>
---	---

<p>cache entry(ies) name cache entries cache footprint K cache footprint cache locality cache locality benet cache memory cache memory unit cache coherence cache coherence protocol</p>	<p>a cache engine cluster cache engine in a cluster cache size(s) level cache sizes variable cache sizes second level cache sizes impact of cache sizes topology with variable cache sizes</p>
---	---

4.3 Colocados contiguos de las combinaciones especializadas.

PRE +	COMBINACIÓN	+ POST
router's, IP	ARP cache	holds, table(s), times
semi-	automatic cache	conscious
second, single, hig-density	Cisco cache	engine(s)
network-	integrated cache	engine
second, first, multi-	level cache	hierarchies, memory, read, sizes, write
direct, sector-, increased	mapped cache	memory, partitions
NetBIOS,	name cache	acts, entries, parameters
web, large	proxy cache	designed, server
8K-byte. Decoding	SEWD cache	chip, makes
	single cache	clock, engine
harvest, IOS, local, URL-based	web cache	communication, control, offers, redirection, stores
different, increases, L1, modified, physical, same, single, unused, L2, two	cache block(s)	size(s), utilization, transfer, due to
implementing, architectures, cluster-based, directory-based, intra-cluster, MESI, nodes, snoopy, typical, modified	cache coherence	protocol(s), schemes, problem, enforced,
apply, applying, semi-automated, semi-automatic, best-case, implementing, natural, performs, results, transparent	cache-conscious	design, heap, data, tree(s), techniques, memory, speedup, layout, counterpart, pointer-based, clustering, functions, version
max-age, powerful, sophisticated, use, web	cache control	setting, mechanisms, request-headers, protocol
certain, Cisco, determines, each, network-integrated, one, less, new, overloaded, redundant, failed, third, additional, single, bypass, cluster, hot-insert, multiple, place, station, name, remaining, three, high-speed	cache engine(s)	cluster, does, 590, transparently, deterministically, 500, connected, products, availability, detects, homes, content, determines, act, assigns, solution, operation, becomes, provides, sends, operates, receives, load, ISP, enterprise
name, high-speed, same	cache entry(ies)	come, consists, stays
4K, 8K, reasonable	cache footprint	
data, improving, increase, instruction, potential, structure's, improve	cache locality	benefit, pointer-based
level 2, fastest, identifying, level, sector-mapped, instruction	cache memory	unit(s), address, provides, effects, uses, simulation, hierarchy
different, name	cache parameters	
analyze, optimized, coloring-improve, estimate, optimizing, program's, improving	cache performance	
optimizes, Pathfire's, LAN, proxy	cache server	stores, load, farms
data, memory, available, level, variable	cache size(s)	required, increase

6. Familia de throughput. Forma analizada: THROUGHPUT

1. FRECUENCIA.

KEYWORD	Frec. Telec	Frec. Lacell	Ratio	Término, Chung	KEYNESS	P valor
153. THROUGHPUT	806	20	153,054305	ESPECIFICO	2.349,00	0
Miembros relacionados:	F Telec	F Gene	Ratio	Término	Keyness	P valor
THROUGHPUTS	14	0	inf/esp	inf/esp	43,9	0
Familia técnica: Sí, todos los miembros son valorados como específicos.						

2. DISTRIBUCIÓN.

Distribución por áreas	Distribución Keyness	Keykeyword	Nº de textos	Porcentaje
9	082	THROUGHPUT	60	3,63%
<i>Dispersion plot</i>				
Area	Words	Hits	Per 1,000	
082 Esp. Tele	997.727	218	0.22	
4 Signal proc.	580.936	120	0.21	
7 Systems	307.655	63	0.20	
081 Esp. Sign	867.208	146	0.17	
3 Telematics	1.205.064	183	0.15	
2 Ar. Comp	329.643	36	0.11	
1 Electronics	722.823	38	0.05	
5 Materials	101.241	1	0.01	
6 Business	373.079	2	0.01	

3. COLOCACIÓN.

3.1 Colocados.

N	Collocates	TOTAL	LEFT	RIGHT	L5	L4	L3	L2	L1	*	R1	R2	R3	R4	R5
1	THE	630	367	263	44	41	51	105	126	0	2	94	52	58	57
2	OF	265	106	159	28	20	21	23	14	0	90	15	16	21	17
3	AND	258	128	130	16	14	28	28	42	0	45	26	15	28	16
4	TO	206	124	82	19	32	48	22	3	0	15	10	14	21	22
5	A	160	94	66	25	12	12	29	16	0	0	32	8	7	19
6	IS	143	44	99	22	14	4	4	0	0	32	10	20	19	18
7	IN	129	64	65	14	11	10	14	15	0	22	14	6	19	4
8	FOR	127	57	70	10	5	19	14	9	0	26	11	11	11	11
9	AS	78	31	47	6	8	8	4	5	0	10	15	7	7	8
10	THAT	75	48	27	9	10	17	11	1	0	5	5	4	8	5
11	WITH	61	27	34	4	4	7	8	4	0	10	7	5	4	8
12	THIS	56	19	37	7	4	5	0	3	0	0	11	13	8	5
13	DATA	55	43	12	4	2	4	1	32	0	2	1	2	3	4
14	BY	52	15	37	5	7	2	1	0	0	9	8	12	5	3
15	LATENCY	51	27	24	2	4	5	8	8	0	3	10	7	3	1
16	CAN	50	28	22	2	12	11	3	0	0	6	2	4	6	4
17	HIGH	50	44	6	2	0	3	3	36	0	0	2	4	0	0
18	SYSTEM	46	34	12	5	5	4	4	16	0	0	0	8	3	1
19	RATE	45	32	13	2	9	4	15	2	0	7	1	0	1	4
20	NETWORK	41	29	12	7	3	3	0	16	0	0	1	3	2	6

21	THAN	38	13	25	4	3	2	4	0	0	13	1	5	3	3
22	ON	36	17	19	6	1	3	4	3	0	4	2	4	4	5
23	ARE	35	9	26	3	4	1	1	0	0	3	4	4	7	8
24	AT	35	22	13	5	9	7	1	0	0	7	2	0	3	1
25	OR	35	16	19	4	3	1	5	3	0	2	5	3	7	2
26	INCREASE	34	25	9	2	3	5	10	5	0	2	2	2	2	1
27	BE	33	10	23	5	4	1	0	0	0	0	13	2	4	4
28	MAXIMUM	33	32	1	0	0	1	4	27	0	0	0	1	0	0
29	WE	32	18	14	8	6	4	0	0	0	2	4	3	1	4
30	DELAY	30	17	13	2	3	2	7	3	0	0	5	3	2	3
31	FIGURE	30	20	10	5	5	7	3	0	0	1	3	0	2	4
32	TIME	30	17	13	6	1	3	4	3	0	0	0	6	3	4
33	IT	29	14	15	5	6	2	1	0	0	0	5	0	7	3
34	AN	28	7	21	1	0	3	3	0	0	0	12	4	2	3
35	AVERAGE	27	19	8	0	1	1	1	16	0	1	4	1	2	0

3.2 Colocados significativos.

Nº	Colocados	Z	T	MI	Nº	Colocados	Z	T	MI
1	SLEEPS	30,64	2,43	7,31	27	REPRESENTATIVE	7,96	2,08	3,87
2	SDSL	39,64	4,31	6,40	28	DECREASES	7,94	2,08	3,86
3	MAXIMIZE	35,37	3,95	6,33	29	PERIODIC	9,08	2,45	3,78
4	MAXIMIZATION	19,06	2,21	6,22	30	METRICS	9,38	2,61	3,69
5	SCREENING	18,85	2,21	6,19	31	DECREASE	7,27	2,06	3,64
6	DEGRADE	20,43	2,60	5,95	32	UTILIZATION	7,03	2,05	3,56
7	SATISFIES	22,28	3,10	5,69	33	RELIABILITY	9,21	2,86	3,38
8	KBPS	28,19	4,15	5,53	34	PEAK	7,35	2,37	3,26
9	LATENCY	46,88	6,98	5,49	35	MODEM	8,73	2,83	3,25
10	AGGREGATE	21,26	3,51	5,20	36	SENDER	7,74	2,53	3,23
11	ACHIEVES	13,75	2,38	5,07	37	THREADS	6,67	2,19	3,22
12	OPTIMIZE	16,01	2,90	4,93	38	OPTIMIZATION	7,42	2,51	3,13
13	CONSTRAINED	16,51	3,05	4,87	39	TRADE	6,34	2,16	3,10
14	EXCEEDS	12,74	2,37	4,86	40	CONSUMPTION	6,17	2,15	3,04
15	SCHEDULE	22,09	4,20	4,79	41	NYSE	12,28	2,36	4,76
16	CONSTRAINT	22,27	4,31	4,74	42	INDOOR	19,15	3,73	4,72
17	OBSERVATIONS	13,40	2,71	4,61	43	FEDEX	11,67	2,35	4,62
18	ABR	9,98	2,13	4,45	44	PAGER	14,21	2,88	4,61
19	OC	9,71	2,13	4,38	45	NASDAQ	15,86	3,31	4,52
20	IMPROVEMENTS	12,72	2,85	4,32	46	UWB	10,16	2,14	4,50
21	DUPLEX	10,19	2,32	4,27	47	BLUETOOTH	39,74	8,48	4,46
22	MOTION	12,15	2,84	4,20	48	WTLS	9,98	2,13	4,45
23	TRANSIENT	9,62	2,31	4,12	49	BLACKBERRY	14,70	3,16	4,43
24	DEGRADATION	10,26	2,49	4,09	50	DECT	13,99	3,02	4,43
25	MBPS	13,41	3,26	4,08	51	YAHOO	10,50	2,33	4,35
26	PLACEMENT	10,10	2,49	4,05	52	OUTDOOR	14,04	3,15	4,31

3.3 Definición del nodo en el diccionario.

- Definición de *throughput* en el *Advanced Learner's Dictionary* de Cambridge (2003):

Throughput (noun): *an amount of work, etc. done in a particular period of time: We need to improve our throughput because demand is high at present.*

- Definición de *throughput* en el diccionario especializado Webster:

Throughput (noun):

- Output relative to input; the amount passing through a system from input to output (especially of a computer program over a period of time).*

Specialty Definition:

In computing:

1. The rate at which a processor can work expressed in instructions per second or jobs per hour or some other unit of performance.
2. Data transfer rate.

3.4 Colocados contiguos.

a) Colocado + <i>throughput</i> .		
Total colocados seleccionados: 15	Frecuencia máxima: 32	Frecuencia mínima: 4
<ul style="list-style-type: none"> - Sustantivos, total 3: <i>data 32, network 16, system 16</i>. - Adjetivos, total 7: <i>high 32/higher 13, maximum 25, average 15, actual 8, overall 7, better 4</i>. - Participios verbales, total 2: <i>given 7, reduced 4</i>. - Formas personales de verbos, total 2: <i>aggregate 8, improve 4</i>. - Siglas, total 1: <i>SDSL 9</i>. 		
b) <i>Throughput</i> + colocado.		
Total colocados seleccionados: 14	Frecuencia máxima: 20	Frecuencia mínima: 3
<ul style="list-style-type: none"> - Sustantivos, total 9: <i>constraint 20, optimization 7, values 7, rates 5, degradation 4, requirements 4, stability 4, maximization 3, value3</i>. - Participios verbales, total 5: <i>achieved 7, constrained 6, screening 4, measured 3, using 3</i>. 		

3.5 Identificación de colocados contiguos especializados en las líneas de concordancia.

*samples between each ISR. The frame-based **throughput rate** is therefore many times higher and type of I/O ports, services, and overall **throughput rate** and features. The value of C structure in Refs. [9, 10 and 11] increase the **throughput rate** of the digit-serial architecture offered by these structures will increase the **throughput rate** of the digit-serial architecture on-processing rate of less than 1000 tps. The **throughput rate** of the IP SCP can support mis designed to achieve an effective transaction-**throughput rate** of 40,000 transactions per second of 800 kHz will impact the return path data **throughput rate**. From the business case variants in modem technology that have allowed **throughput rates** to grow to 28.8 Kbps, with compression standards enabling effective **throughput rates** for some kinds of communication frame-based processing achieves higher **throughput rates** Continuous-Time Signals. Most real-time DSP systems optimize **throughput rates** by processing data in "batches". GPRS will offer a tenfold increase in data **throughput rates**, from 9.6kbit/s to 115kbit/s.*

Concordancias de *throughput rate(s)*.

a) Colocado + *throughput*:

<i>actual throughput</i>	<i>high(er) throughput</i>
aggregate throughput	<i>maximum throughput</i>
<i>average throughput</i>	<i>network throughput</i>
<i>data throughput</i>	<i>overall throughput</i>
<i>system throughput</i>	<i>wireless throughput</i>

b) *Throughput* + colocado:

<i>throughput constrained</i>	<i>throughput optimization</i>
<i>throughput degradation</i>	<i>throughput rates</i>
<i>throughput maximization</i>	<i>throughput requirements</i>
<i>throughput measured</i>	<i>throughput screening</i>
<i>throughput stability</i>	<i>throughput value(s)</i>

4. AGRUPACIONES LÉXICAS O CLUSTERS.

4.1 Distribución de clusters.

Combinaciones	Volumen	Frecuencia máxima
2-cluster	84	12
3-cluster	98	44
4-cluster	46	23
5-cluster	27	8

4.2 Inclusión de combinaciones especializadas en unidades más amplias.

<i>average throughput</i> <i>aggregate average throughput</i> <i>throughput requirement(s)</i> <i>minimum throughput requirement</i>	<i>high throughput</i> <i>high throughput screening</i> <i>throughput screening</i> <i>high throughput screening</i>
---	---

4.3 Colocados contiguos de las combinaciones especializadas.

PRE +	COMBINACIÓN	+ POST
	<i>actual throughput</i>	exceeds
<i>provide, Ethernet-with, obtaining</i>	aggregate <i>throughput</i>	
aggregate , <i>value</i>	<i>average throughput</i>	decreases , <i>achieved, in function of</i>
<i>increased, using, higher, more, including, increase, yield, maximum, real, path, Mbps, effective, maintain, real-time, control, possible</i>	<i>data throughput</i>	<i>measurements, potential, rate(s), using, constraints</i>
<i>achieving, deliver, require, achieve, generates, discarded, new, making, building, proportionately, permits, offers</i>	<i>high(er) throughput</i>	<i>connections, equipment, requirements, screening, values, choice, communication, component, environments, measurement, placement, switches, systems, transaction, lower, achieved</i>
<i>P-HTTP, deliver, attained, theoretical</i>	<i>maximum throughput</i>	<i>achieved, potential, supported, value, resynchronization</i>
<i>wireless, increasing, include, reduces</i>	<i>network throughput</i>	<i>derived, drops</i>
<i>near-optimal</i>	<i>overall throughput</i>	<i>efficiency, rate</i>
<i>high, pessimistic, improve, overall, available, project, computer</i>	<i>system throughput</i>	<i>requirements</i>
<i>Boost, understand</i>	<i>wireless throughput</i>	
<i>given, pre-specified, data-</i>	<i>throughput constraint(s)</i>	
<i>widespread, concomitant,</i>	<i>throughput degradation</i>	<i>occurs</i>
<i>off-line, MPLS</i>	<i>throughput latency</i>	<i>computation, IP</i>
<i>on-line</i>	<i>throughput optimization</i>	<i>algorithm, experiments</i>
<i>frame-based, overall, transaction-, data, allowed, effective, higher, optimize</i>	<i>throughput rate(s)</i>	
<i>minimum, reduced, systems, high</i>	<i>throughput requirement(s)</i>	
<i>high</i>	<i>throughput screening</i>	<i>instruments</i>
<i>maximum, theoretical, fixed, high</i>	<i>throughput value(s)</i>	<i>pairs, calculated, given</i>

7. Familia de *latency*. Forma analizada: LATENCY

1. FRECUENCIA.

KEYWORD	Frec. Telec	Frec. Lacell	Ratio	Término, Chung	KEYNESS	P valor
151. LATENCY	777	7	421,563967	ESPECIFICO	2.360,30	0
Miembros relacionados:	F Telec	F Gene	Ratio	Término	Keyness	P valor
LATENCIES	52	0	inf/esp	inf/esp	163,1	0
Familia técnica: Sí, todos los miembros son valorados como específicos.						

2. DISTRIBUCIÓN.

Distribución por áreas	Distribución Keyness	Keykeyword	Nº de textos	Porcentaje
9	4,082	LATENCY	51	3,08%
<i>Dispersion plot</i>				
Area	Words	Hits	Per 1,000	
4 Signal proc.	580.936	206	0.35	
2 Ar. Comp	329.643	83	0.25	
082 Esp. Tele	997.727	198	0.20	
7 Systems	307.655	61	0.20	
1 Electronics	722.823	85	0.12	
3 Telematics	1.205.064	86	0.07	
081 Esp. Sign	867.208	55	0.06	
5 Materials	101.241	1	0.01	
6 Business	373.079	1	0.00	

3. COLOCACIÓN.

3.1 Colocados.

N	Collocates	TOTAL	LEFT	RIGHT	L5	L4	L3	L2	L1	*	R1	R2	R3	R4	R5
1	THE	516	307	209	54	55	58	59	81	0	4	61	53	51	40
2	AND	258	114	144	17	20	24	23	30	0	59	28	23	17	17
3	TO	184	114	70	25	20	31	35	3	0	7	19	15	12	17
4	OF	180	90	90	22	11	16	15	26	0	44	7	9	12	18
5	IS	141	36	105	8	14	7	6	1	0	41	11	16	23	14
6	IN	139	51	88	18	12	9	5	7	0	25	15	21	11	16
7	A	129	66	63	9	18	13	19	7	0	2	24	8	21	8
8	FOR	124	37	87	9	4	12	4	8	0	35	19	10	8	15
9	AS	90	39	51	9	12	8	6	4	0	16	7	10	11	7
10	LOW	74	62	12	2	1	4	0	55	0	0	7	3	1	1
11	NETWORK	70	50	20	3	9	9	6	23	0	0	1	7	5	7
12	THAT	70	43	27	8	16	8	11	0	0	4	5	7	5	6
13	WITH	62	36	26	4	8	11	9	4	0	2	10	4	6	4
14	BY	58	17	41	6	1	3	5	2	0	8	9	11	11	2
15	OUTPUT	56	52	4	0	1	3	0	48	0	0	0	1	3	0
16	ON	54	27	27	9	5	2	9	2	0	3	5	11	4	4
17	MEMORY	53	47	6	1	1	3	0	42	0	0	1	1	1	3
18	PERIODIC	52	50	2	0	1	0	49	0	0	0	0	2	0	0
19	THROUGHPUT	51	24	27	1	3	4	9	7	0	2	11	6	5	3
20	TIME	51	21	30	8	3	3	5	2	0	2	3	11	5	9

21	THIS	50	24	26	1	9	7	5	2	0	0	9	7	2	8
22	END	42	39	3	4	6	14	5	10	0	0	1	0	2	0
23	BE	40	13	27	4	6	2	1	0	0	0	9	4	8	6
24	PACKET	36	21	15	7	1	2	2	9	0	1	5	3	2	4
25	CAN	34	9	25	3	3	2	1	0	0	6	1	7	4	7
26	ARE	33	14	19	5	5	2	0	2	0	1	2	5	6	5
27	JITTER	32	5	27	0	0	1	4	0	0	0	18	7	1	1
28	COMMUNICATION	30	15	15	4	2	1	1	7	0	6	1	1	4	3
29	OVERALL	30	24	6	0	1	4	13	6	0	0	0	4	2	0
30	BOUND	29	7	22	3	2	2	0	0	0	19	2	0	0	1
31	FROM	28	7	21	1	2	1	1	2	0	4	4	3	3	7
32	MAXIMUM	28	16	12	0	0	0	13	3	0	0	0	3	5	4
33	SYSTEM	27	5	22	3	1	1	0	0	0	1	2	10	3	6
34	AN	26	13	13	3	1	1	8	0	0	0	2	2	4	5
35	DELAY	26	12	14	2	2	2	5	1	0	2	4	3	4	1

3.2 Colocados significativos.

Nº	Colocados	Z	T	MI	Nº	Colocados	Z	T	MI
1	MRR	41,60	2,98	7,60	23	REDUCES	14,78	3,29	4,33
2	TASKING	45,02	3,44	7,42	24	METRICS	13,47	3,14	4,20
3	RESYNCHRONIZATION	27,64	2,22	7,27	25	REDUCING	14,04	3,28	4,20
4	JITTER	62,28	5,61	6,94	26	VARIES	9,06	2,11	4,20
5	PERIODIC	73,37	7,14	6,72	27	EXPERIENCED	9,85	2,31	4,18
6	INTERRUPT	44,20	4,84	6,38	28	HANDOVER	9,47	2,30	4,08
7	GUARANTEED	40,40	5,02	6,02	29	GUARANTEES	8,09	2,09	3,91
8	TRIP	24,37	3,11	5,94	30	SIZES	9,84	2,63	3,81
9	BOUNDS	22,45	3,10	5,71	31	CONSTRAINT	10,93	2,93	3,80
10	CONSTRAINED	27,28	3,92	5,60	32	CYCLES	9,08	2,45	3,78
11	BOUND	36,41	5,27	5,58	33	REDUCTION	12,32	3,34	3,77
12	ALGORITHMIC	16,24	2,40	5,52	34	SCHEDULING	12,30	3,34	3,76
13	MILLISECONDS	13,87	2,18	5,34	35	MINIMIZE	8,30	2,27	3,74
14	QUEUEING	18,10	2,92	5,26	36	MESH	7,46	2,06	3,71
15	MS	26,41	4,45	5,14	37	SCHEDULE	9,01	2,59	3,59
16	FAIRNESS	13,54	2,55	4,82	38	MINIMAL	6,70	2,03	3,44
17	CLOCKS	11,32	2,16	4,79	39	TRADE	7,71	2,39	3,38
18	ROUND	18,24	3,47	4,78	40	CYCLE	9,14	2,85	3,36
19	CONTENTION	11,09	2,15	4,73	41	CUT	6,40	2,01	3,33
20	LOWEST	14,87	3,03	4,59	42	STREAM	7,69	2,65	3,08
21	TRANSIENT	13,29	2,71	4,59	43	COMPUTATION	7,03	2,48	3,01
22	THREAD	16,31	3,56	4,39					

3.3 Definición del nodo en el diccionario.

- Definición de *Latency* en el *Advanced Learner's Dictionary* de Cambridge (2003):

Latency (noun): (formal) at latent. **Latent** (adj): sent but needing particular conditions to become active, obvious or completely developed: Recent developments in the area have brought latent ethnic tension out into the open.

- Definición de *latency* en el diccionario especializado Webster:

Latency (noun):

- The time it takes for a packet to cross a network connection, from sender to receiver.
- The period of time that a frame is held by a network device before it is forwarded.

Specialty Definition:

Latency: The engineering definition of latency is the time a message takes to traverse a system. Latency is closely tied to another engineering concept, throughput. While each is a measure of

4. AGRUPACIONES LÉXICAS O CLUSTERS.

4.1 Distribución de clusters.

Combinaciones	Volumen	Frecuencia máxima
2-cluster	80	99
3-cluster	21	92
4-cluster	3	16

4.2 Inclusión de combinaciones especializadas en unidades más amplias.

<p>interrupt latency maximum interrupt latency</p> <p>memory latency cluster memory latency network memory latency overall memory latency overall cluster memory latency overall network memory latency</p>	<p>output latency output latency value</p> <p>switch latency thread switch latency trip latency round trip latency guarantees round trip latency</p> <p>latency bounds lower latency bounds</p>
---	--

4.3 Colocados contiguos de las combinaciones especializadas.

PRE +	COMBINACIÓN	+ POST
<i>creates</i>	<i>additional latency</i>	<i>added, due to</i>
<i>vertical, total, high, large</i>	handover latency	
<i>typical, very, fairly, relatively, support</i>	<i>high(er) latency</i>	<i>associated, factor, inherent, environments, bound, connections</i>
<i>maximum, minimum, lowest, reduce</i>	interrupt latency	<i>period, related, varies</i>
<i>ultra, consistently, provides, providing, delivers, achieve, very, requires, supports, achieving, provide, simple, demand, implements, non-,</i>	<i>low(er) latency</i>	<i>allows, bound(s), kernel, communication(s), first, requirements, support, array, acquisition, class, constructs, handoffs, highly, one-sided, queue, queuing, schedule, service, small, system, techniques, traffic</i>
<i>overall, network, applications, lowest, cluster, manifestation, investigates, impact on, tolerate, long, total</i>	<i>memory latency</i>	<i>component, decreased, decreases, followed, time</i>
<i>increases, reducing, due to, amount of, considerable, impact of, one-way, minimize, associated</i>	<i>network latency</i>	
periodic-	<i>output latency</i>	computation , <i>depends, gives, increases, metric, needs, relates, significantly, using, value</i>
<i>end-to-end, data, control</i>	<i>packet latency</i>	
thread , <i>per-</i>	<i>switch latency</i>	<i>dep, depends, divided, guaranteed</i>
<i>called, delay</i>	tasking latency	
round-	trip latency	<i>budget, depending, guarantee</i>
<i>analytical, least, lowest, start-up, higher, lower, guarantee, round-trip</i>	<i>latency bound(s)</i>	<i>calculation</i>
<i>low, lower, lowest,</i>	<i>latency communication(s)</i>	<i>operations, library, service</i>
<i>off-line, periodic-output, throughput-</i>	<i>latency computation</i>	<i>method, scheme</i>
<i>resource-and-, optimal</i>	<i>latency-constrained</i>	scheduling , resynchronization(s)
<i>Improved, new, consequent</i>	<i>latency-reduction</i>	<i>scheme, technique, post-processing</i>

8. Familia de impedance. Forma analizada: IMPEDANCE

1. FRECUENCIA.

KEYWORD	Frec. Telec	Frec. Lacell	Ratio	Término, Chung	KEYNESS	P valor
225. IMPEDANCE	576	13	168,275014	ESPECIFICO	1.687,80	0
Miembros relacionados:	F Telec	F Gene	Ratio	Término	Keyness	P valor
IMPEDANCES	55	3	69,6276822	ESPECIFICO	150,3	0
TRANSIMPEDANCE	30	0	inf/esp	inf/esp	94,1	0
Familia técnica: Sí, todos los miembros son valorados como específicos.						

2. DISTRIBUCIÓN.

Distribución por áreas	Distribución Keyness	Keykeyword	Nº de textos	Porcentaje
8 (- 6)	1, 4	IMPEDANCE	47	2,84%
<i>Dispersion plot</i>				
Area	Words	Hits	Per 1,000	
1 Electronics	722.823	283	0.39	
4 Signal proc.	580.936	155	0.27	
5 Materials	101.241	11	0.11	
081 Esp. Sign	867.208	79	0.09	
2 Ar. Comp	329.643	8	0.02	
3 Telematics	1.205.064	24	0.02	
082 Esp.Tele	997.727	15	0.02	
7 Systems	307.691	1	0.00	

3. COLOCACIÓN.

3.1 Colocados.

N	Collocates	TOTAL	LEFT	RIGHT	L5	L4	L3	L2	L1	*	R1	R2	R3	R4	R5
1	THE	522	295	227	45	43	37	113	57	0	2	77	62	36	76
2	OF	214	86	128	16	12	23	27	8	0	65	18	8	20	30
3	AND	170	68	102	10	11	18	22	7	0	34	27	8	14	7
4	A	143	82	61	14	14	13	41	0	0	1	25	11	13	36
5	TO	140	83	57	22	27	20	12	2	0	8	14	11	12	31
6	INPUT	135	118	17	4	6	7	1	100	0	2	3	0	8	37
7	IS	127	40	87	15	8	9	8	0	0	31	7	14	19	27
8	IN	69	29	40	7	4	12	2	4	0	9	8	4	9	19
9	OUTPUT	66	49	17	1	3	1	0	44	0	0	4	8	4	13
10	HIGH	64	57	7	1	1	2	13	40	0	0	3	2	1	4
11	AN	54	40	14	2	4	6	7	21	0	0	4	5	0	18
12	FOR	54	20	34	3	6	4	6	1	0	4	6	6	8	1
13	WITH	51	33	18	10	5	11	7	0	0	3	1	7	5	4
14	LOW	48	38	10	2	1	2	12	21	0	2	2	2	3	7
15	AS	43	21	22	4	5	7	4	1	0	2	3	10	4	10
16	BE	43	11	32	4	4	2	1	0	0	0	10	4	9	11
17	CHARACTERISTIC	43	41	2	0	1	0	0	40	0	0	1	0	1	9
18	BY	42	23	19	3	8	4	5	3	0	0	7	7	3	3
19	THIS	38	12	26	3	3	2	0	4	0	0	9	8	3	4
20	ARE	36	10	26	3	4	1	1	1	0	3	9	3	4	10

21	LINE	33	15	18	1	3	5	4	2	0	3	3	2	5	2
22	THAT	29	17	12	3	2	6	6	0	0	2	4	3	2	5
23	OR	28	14	14	4	2	2	6	0	0	10	1	1	0	8
24	AT	27	5	22	2	2	1	0	0	0	9	6	1	3	3
25	CAN	27	2	25	1	1	0	0	0	0	3	3	10	6	0
26	LOOP	27	18	9	2	1	1	13	1	0	0	0	2	3	4
27	MATCHING	27	1	26	0	1	0	0	0	0	26	0	0	0	4
28	LOAD	26	19	7	1	0	3	0	15	0	1	0	2	3	2
29	Z	24	4	20	1	2	0	1	0	0	8	6	3	3	4
30	ADMITTANCE	22	5	17	1	3	0	1	0	0	1	13	0	3	6
31	BANDWIDTH	22	6	16	2	3	0	1	0	0	12	2	1	0	3
32	CIRCUIT	22	9	13	2	2	1	2	2	0	0	3	3	6	1
33	CURRENT	22	8	14	4	2	2	0	0	0	0	2	2	5	5
34	EQUIVALENT	21	20	1	1	2	5	7	5	0	0	0	0	0	2
35	GAIN	20	12	8	5	2	1	3	1	0	0	2	4	0	6

3.2 Colocados significativos.

Nº	Colocados	Z	T	MI
1	ZS	88,50	2,64	10,13
2	ADMITTANCE	70,02	4,67	7,81
3	LOCUS	28,60	2,22	7,37
4	TRANS	24,30	2,22	6,91
5	OHMS	28,49	2,62	6,88
6	CHARACTERISTIC	68,02	6,50	6,78
7	SMITH	25,09	3,11	6,02
8	MATCHED	30,17	3,81	5,97
9	CHART	26,98	3,41	5,97
10	MATCHING	37,74	5,10	5,77
11	MISMATCH	17,56	2,40	5,74
12	INDUCTANCE	14,30	2,18	5,42
13	TRANSFORMER	14,17	2,18	5,40
14	INVERTING	18,26	3,07	5,14
15	INFINITE	17,59	3,07	5,04
16	FEED	15,65	2,74	5,03
17	LOADED	13,54	2,71	4,64
18	MATCH	17,63	3,59	4,59
19	AMP	10,85	2,34	4,43
20	CLOSED	14,42	3,16	4,38
21	PLOT	10,35	2,33	4,31
22	AMPLIFIERS	9,44	2,30	4,07
23	REFLECTION	8,57	2,10	4,05
24	BOUNDARY	8,17	2,09	3,94
25	CONDITION	10,86	2,80	3,91
26	TERMINALS	8,83	2,44	3,71
27	OP	7,39	2,06	3,68
28	CAPACITOR	7,31	2,06	3,66

3.3 Definición del nodo en el diccionario.

1. El *Advanced Learner's Dictionary* de Cambridge (2003) no registra la definición de *impedance*.
2. Definición de *impedance* en el diccionario especializado Webster:

Impedance (Noun): A material's opposition to the flow of electric current; measured in ohms.

Specialty Definition:

In electrical engineering:

The apparent opposition in an electrical current to the flow of an alternating current that is analogous to the actual electrical resistance to a direct current and that is the ratio of electromotive force to the effective current.

In computing:

The total opposition offered by a component or circuit to the flow of an alternating or varying current: a combination of resistance and reactance.

3.4 Colocados contiguos.

a) Colocado + <i>impedance</i> .		
Total colocados seleccionados: 43	Frecuencia máxima: 75	Frecuencia mínima: 3
<ul style="list-style-type: none"> - Sustantivos, total 8: <i>input 100, output 44, load 15, source 7, broadband 5, top 5, port 4, system 4.</i> - Adjetivos, total 5: <i>characteristic 40, high 40, low 40, good 7, internal 6.</i> 		
b) <i>Impedance</i> + colocado.		
Total colocados seleccionados: 18	Frecuencia máxima: 26	Frecuencia mínima: 3
<ul style="list-style-type: none"> - Sustantivos, total 13: <i>bandwidth 12, state 8, locus 5, nature 5, boundary 4, condition 4, mismatch 4, values 4, levels 3, measurements 3, mode 3, point 3, Smith 3.</i> - Participios verbales, total 3: <i>matching 26, matched 8, seen 5.</i> - Formas personales de verbo, total 1: <i>match 12.</i> - Siglas, total 1: <i>Z 8.</i> 		

3.5 Identificación de colocados contiguos especializados en las líneas de concordancia.

sity modulation bandwidth, modulation efficiency and impedance matching of semiconductor cascade las to an arbitrary number of devices and provide better impedance matching and enhanced RF gain over wicitation of parallel devices, and to achieve broadband impedance matching concurrent with high-power am aAs. A uniform 12 GHz bandwidth was achieved by impedance matching the entire array to a 50 load.

Impedance Matching and the Smith Chart: The Fund

it is possible to use this structure to obtain excellent impedance matching without jeopardising the intensi ccommodate small tolerance variations in the filter. Impedance Matching Filters and Diplexers. Becaus publication 1 Microwave Filters, Impedance Matching Networks and Coupling Struct

ould therefore be made no larger than is required for impedance matching slot width: The width of the slo plications of antennas and antenna arrays with good impedance matching and VSWR characteristics ov loading by shaped slot, notch, cuts, pin or post. iii. Impedance matching of the feed: Probe compensati e is measured with an inductor of $Q = 40$ in the input impedance matching circuit. Note 3. Specs are bas n of systems using transistors, complex discrete LC impedance matching networks and transmission line . A theoretical study shows that excellent microwave impedance matching over more than an octave band % impedance bandwidth. Another new technique of impedance matching by capacitive loading of invert PSNR is the primary consideration, the principles of impedance matching are identical to the load-match an effective load capacitance of 34 fF would produce impedance matching to 50 . From the models discu sign a network to insert between them so that proper impedance matching occurs. Figure 11. The repres the Smith Chart: The Fundamentals. Tutorial on RF impedance matching using the Smith Chart. Exampl nclusion, a novel travelling wave configuration for the impedance matching of laser diode arrays has been as inductors, capacitors, and resistors is followed by impedance matching, tuning, resonators, and filters. rget component values. There are many ways to do impedance matching, including: Computer simulatio ns to introduce multiple resonances as well as input impedance matching, (ii) Genetic Algorithm (GA) ba function of the Smith chart is the ability to determine impedance matching. This is the reverse operation o pol, improved structural stability, and improved input impedance matching. REFERENCES.

are dedicated to differing design functions and not to impedance matching. Designers have to be familiar

Concordancias de *impedance matching*.

a) Colocado + impedance:

<i>broadband impedance</i>	<i>low impedance</i>
characteristic impedance	<i>output impedance</i>
<i>high impedance</i>	<i>port impedance</i>
<i>input impedance</i>	<i>source impedance</i>
<i>internal impedance</i>	<i>system impedance</i>
<i>load impedance</i>	<i>top impedance</i>

b) Impedance + colocado:

<i>impedance bandwidth(s)</i>	<i>impedance matched</i>	<i>impedance point</i>
<i>impedance boundary</i>	<i>impedance matching</i>	<i>impedance Smith</i>
<i>impedance condition</i>	<i>impedance measurements</i>	<i>impedance state</i>
<i>impedance levels</i>	<i>impedance mismatch</i>	<i>impedance values</i>
<i>impedance locus</i>	<i>impedance mode</i>	<i>impedance Z</i>
<i>impedance match</i>	<i>impedance nature</i>	

4. AGRUPACIONES LÉXICAS O CLUSTERS.

4.1 Distribución de clusters.

Combinaciones	Volumen	Frecuencia máxima
2-cluster	73	11
3-cluster	80	36
4-cluster	42	14
5-cluster	23	10

4.2 Inclusión de combinaciones especializadas en unidades más amplias.

<p><i>characteristic impedance</i> <i>characteristic impedance values</i> <i>line characteristic impedance values</i> <i>transmission line characteristic impedance values</i></p> <p><i>input impedance</i> <i>input impedance nature</i> <i>input impedance matching</i> <i>high input impedance nature</i></p> <p><i>load impedance</i> <i>load impedance down</i> <i>equivalent load impedance down</i></p> <p><i>impedance values</i> <i>characteristic impedance values</i> <i>line characteristic impedance values</i> <i>transmission line characteristic impedance values</i></p>	<p><i>impedance boundary</i> <i>impedance boundary condition</i> <i>impedance condition</i> <i>high-impedance condition</i> <i>impedance match</i> <i>input impedance matching</i> <i>impedance nature</i> <i>input impedance nature</i> <i>high input impedance nature</i> <i>impedance Smith</i> <i>impedance Smith chart</i> <i>impedance state</i> <i>high-impedance state</i></p>
---	--

4.3 Colocados contiguos de las combinaciones especializadas.

PRE +	COMBINACIÓN	+ POST
<i>excellent, achieve</i>	<i>broadband impedance</i>	<i>match, -matched, matching</i>
<i>quasi-static, complex, effective, specified, system, Z, non-loaded</i>	<i>characteristic impedance</i>	<i>determined, deviate, given, reduces, using, valued, ZL</i>
<i>feeding, voltage, requiring, old-fashioned, sensitive,</i>	<i>high impedance</i>	<i>antenna, CM, input, interconnect, interfaces, line, load, sources, state, transmission, Z, voltmeter, condition, earphones, node(s)</i>
<i>high, low, desired, infinite, including, large, current, buffer's, signal, stage, antenna, loop, ohms, improved, calculated, equivalent</i>	<i>input impedance</i>	<i>allows, bandwidth(s), determined, due to, expressed, locus, match, matching, measurements, multiplied, nature, needed Zin, requirements, Rin, summing, ACLin,</i>
<i>effective</i>	<i>internal impedance</i>	<i>approach, EII, equal,, seen</i>
<i>equivalent, anode, high, external</i>	<i>load impedance</i>	<i>presented, values</i>

<i>allows, industry-standard</i>	<i>low impedance</i>	condition , <i>consisting, input, larger, looking, modest, nodes, point, closed, XLR</i>
<i>enhanced, zero, effective, loop, low, finite, well-controlled, equivalent, signal</i>	<i>output impedance</i>	<i>CM, reduced, levels</i>
	<i>port impedance</i>	<i>modelling</i>
	<i>source impedance</i>	ZS
<i>ohm, microwave, external</i>	<i>system impedance</i>	
<i>equivalent, modal</i>	<i>top impedance</i>	<i>ZLm</i>
<i>good, input, maximum, wider, comparable, narrow, demonstrated, measured, improved</i>	<i>impedance bandwidth(s)</i>	<i>VSWR, ranging</i>
<i>new, standard, surface</i>	<i>impedance boundary</i>	condition, condition-based
<i>low, high-</i>	<i>impedance condition</i>	
<i>output, feedback, loop</i>	<i>impedance levels</i>	
<i>calculated, input</i>	<i>impedance locus(i)</i>	<i>versus</i>
<i>approximated, broad-band, good, broadband, improved, input, better</i>	<i>impedance match</i>	<i>observed</i>
<i>broadband</i>	<i>impedance matched</i>	<i>array, photodetector, MSM, systems</i>
<i>better, broadband, good, input, LC, microwave, produce, proper, RF, do, determine</i>	<i>impedance matching</i>	<i>concurrent, filters, networks, slot, circuit, using</i>
<i>term</i>	<i>impedance mode</i>	<i>used</i>
<i>input</i>	<i>impedance nature</i>	
<i>low</i>	<i>impedance point</i>	
<i>suitable, whole</i>	<i>impedance Smith</i>	chart
<i>high, high-</i>	<i>impedance state</i>	<i>denoted</i>
<i>characteristic, load, different</i>	<i>impedance values</i>	

9. Familia de *bluetooth*. Forma analizada BLUETOOTH

1. FRECUENCIA.

KEYWORD	Frec. Telec	Frec. Lacell	Ratio	Término, Chung	KEYNESS	P valor
256. BLUETOOTH	488	2	926,681153	ESPECIFICO	1.505,50	0
Término técnico.						

2. DISTRIBUCIÓN.

Distribución por áreas	Distribución Keynes	Keykeyword	Nº de textos	Porcentaje
7 (- 5,7)	4, 801, 802	BLUETOOTH	27	1,63%
<i>Dispersion plot</i>				
Area	Words	Hits	Per 1,000	
4 Signal proc.	580.936	112	0.19	
082 Esp. Tele	997.727	191	0.19	
081 Esp. Sign	867.208	135	0.16	
3 Telematics	1.205.064	41	0.03	
2 Ar. Comp	329.643	9	0.03	
1 Electronics	722.823	4	0.01	
6 Business	373.079	1	0.00	

3. COLOCACIÓN.

3.1 Colocados.

N	Collocates	TOTAL	LEFT	RIGHT	L5	L4	L3	L2	L1	*	R1	R2	R3	R4	R5
1	THE	264	167	97	24	33	20	12	78	0	0	18	28	23	28
2	AND	143	68	75	11	9	12	10	26	0	12	15	25	8	15
3	A	134	68	66	13	14	9	2	30	0	1	11	19	20	15
4	TO	97	50	47	12	7	8	15	8	0	1	8	13	10	15
5	OF	96	76	20	13	15	10	20	18	0	0	0	1	7	12
6	IS	93	20	73	5	4	2	7	2	0	12	17	23	10	11
7	FOR	90	57	33	16	6	6	7	22	0	1	8	3	5	16
8	WIRELESS	81	20	61	4	8	0	7	1	0	40	5	8	7	1
9	TECHNOLOGY	73	12	61	0	2	3	5	2	0	36	18	2	5	0
10	IN	64	32	32	10	3	3	10	6	0	0	11	8	6	7
11	WITH	60	40	20	8	2	5	14	11	0	0	7	5	5	3
12	DEVICES	46	17	29	0	2	4	5	6	0	13	10	3	2	1
13	AS	44	30	14	5	5	5	6	9	0	0	2	7	4	1
14	BE	41	17	24	6	7	2	2	0	0	0	9	3	2	10
15	OTHER	36	30	6	4	9	4	3	10	0	0	0	1	3	2
16	THIS	33	6	27	2	1	2	0	1	0	2	8	10	1	6
17	THAT	31	21	10	3	3	7	6	2	0	0	3	1	1	5
18	COMMUNICATION	28	14	14	1	0	10	2	1	0	4	9	0	0	1
19	PHONE	28	5	23	0	2	1	2	0	0	5	10	1	1	6
20	TECHNOLOGIES	27	16	11	3	3	5	2	3	0	5	1	1	2	2
21	RADIO	26	9	17	2	2	2	3	0	0	13	0	1	1	2
22	CAN	24	2	22	1	0	1	0	0	0	4	6	3	7	2
23	DATA	24	18	6	3	10	3	2	0	0	0	1	2	3	0
24	OR	24	9	15	3	1	1	1	3	0	5	5	2	3	0

25	WILL	24	10	14	3	3	0	1	3	0	5	1	2	2	4
26	SPECIFICATION	23	5	18	2	1	0	2	0	0	15	1	0	1	1
27	ENABLED	22	9	13	2	1	0	6	0	0	13	0	0	0	0
28	ARE	21	7	14	2	4	1	0	0	0	2	5	3	2	2
29	COMMUNICATIONS	21	8	13	1	2	2	3	0	0	3	10	0	0	0
30	IT	21	8	13	2	0	2	0	4	0	0	2	5	3	3
31	ON	21	9	12	1	2	2	4	0	0	1	4	4	1	2
32	SUCH	21	19	2	1	5	3	9	1	0	0	1	1	0	0
33	BY	20	14	6	1	4	1	6	2	0	0	2	2	1	1
34	MOBILE	20	7	13	2	1	4	0	0	0	6	1	0	6	0
35	HOMERF	19	9	10	1	0	0	6	2	0	0	6	2	2	0

3.2 Colocados significativos.

Nº	Colocados	Z	T	MI
1	CHIPS	7,92	2,08	3,85
2	OPTIMIZED	9,16	2,30	3,99
3	UTILIZATION	9,36	2,12	4,28
4	ENABLING	11,10	2,51	4,29
5	GPS	12,14	2,53	4,52
6	EDITION	13,44	2,37	5,00
7	REPLACEMENT	13,24	2,18	5,21
8	RADIOS	13,38	2,18	5,24
9	FI	19,13	2,77	5,58
10	SOC	15,38	2,19	5,62
11	HANDHELD	20,10	2,60	5,90
12	BASEBAND	21,78	2,78	5,94
13	WI	23,02	2,79	6,09
14	ADAPTER	26,59	3,12	6,18
15	WHITEBOARD	29,59	2,43	7,21
16	WPANS	35,76	2,23	8,01
17	SIG	69,69	4,11	8,17
18	IRDA	50,75	2,82	8,34
19	QUALIFICATION	78,02	4,23	8,41
20	HOMERF	81,69	4,35	8,46
21	BLU	41,97	2,23	8,47
22	SDK	48,50	2,23	8,88
23	SCATTERNETS	48,50	2,23	8,88
24	WAVELAN	91,14	3,60	9,32
25	BQP	71,89	2,82	9,34
26	SCATTERNET	85,10	2,83	9,83
27	BLUETOOTH'S	90,29	2,45	10,41

3.3 Definición del nodo en el diccionario.

1. El *Advanced Learner's Dictionary* de Cambridge (2003) no registra la definición de *bluetooth*.
2. Definición de *bluetooth* en el diccionario especializado Webster:

Bluetooth Specialty Definition:

Bluetooth is an Industrial Specification for Wireless PANs first developed by Ericsson, later formalized by the Bluetooth SIG, which was formally announced 20th May 1999. It was composed by Sony Ericsson, IBM, Intel, Nokia and Toshiba.

It brings together telecoms and IT manufacturers in the creation of a short-range wideband radio standard to allow mobile phones, palmtops and portable PCs to communicate with each other without cables.

3.4 Colocados contiguos.

a) Colocado + <i>bluetooth</i> .		
Total colocados seleccionados: 6	Frecuencia máxima: 3	Frecuencia mínima: 3
<ul style="list-style-type: none"> - Sustantivos, total 2: <i>book</i> 3, <i>networks</i> 3. - Adjetivos, total 1: <i>different</i> 3. - Participios verbales, total 2: <i>integrated</i> 6, <i>called</i> 3. - Siglas, total 1: <i>HTTP</i> 3. 		
b) <i>Bluetooth</i> + colocado.		
Total colocados seleccionados: 25	Frecuencia máxima: 40	Frecuencia mínima: 3
<ul style="list-style-type: none"> - Sustantivos, total 19: <i>wireless</i> 40, <i>technology</i> 36/<i>technologies</i> 5, <i>specification</i> 15, <i>radio</i> 13, <i>qualification</i> 13, <i>devices</i> 13/<i>device</i> 6, <i>adapter</i> 9, <i>scatternet</i> 6/<i>scatternets</i> 4, <i>products</i> 6, <i>connection</i> 5, <i>radios</i> 4, <i>links</i> 4, <i>edition</i> 4, <i>system</i> 3, <i>packet</i> 3, <i>chips</i> 3. - Participios verbales, total 1: <i>enabling</i> 3. - Formas personales de verbos, total 1: <i>enabled</i> 13. - Adjetivos, total 2: <i>mobile</i> 6, <i>cellular</i> 3. - Siglas, total 2: <i>SIG</i> 7, <i>SOC</i> 5. 		

3.5 Identificación de colocados contiguos especializados en las líneas de concordancia.

Fig. 1 shows an example of bandwidth usage of a Bluetooth scatternet before finding a QoS route. The divi 3.1. Phase I: Free Time-Slot Information Collection. A Bluetooth scatternet is assumed to initially be formed Source node S initiates and floods a BQ packet into a Bluetooth scatternet, where D is a destination node, BR The value of the current TTL. Node e floods BQ into the Bluetooth scatternet, where itself FT is the free time-slo rrent FTSL, BR, current TTL) packet from node e_ in the Bluetooth scatternet, the current TTL and D ADDR field denote a free time-slot set for a Bluetooth device in a Bluetooth scatternet. For instance as shown in Fig. 1, { the QoS REQuest, or BQ REQ, packet and floods into Bluetooth scatternets until the BQ REQ packets arrive levels. This work improves the bandwidth utilization of Bluetooth scatternets by providing a new interpiconet s mpted to develop QoS-extension routing scheduling in Bluetooth scatternets. First, Cordeiro et al. [4] propose calculate free time slots between two adjacent nodes in Bluetooth scatternets. Let {á1, á2, ..., ák} denote a fre This work can improve the bandwidth utilization of Bluetooth scatternets. Interpiconet scheduling problems

Concordancias de *Bluetooth scatternet(s)*.

a) Colocado + *bluetooth*:

integrated Bluetooth

b) *Bluetooth* + colocado:

<i>Bluetooth adapter</i>	<i>Bluetooth mobile</i>	<i>Bluetooth SOC</i>
<i>Bluetooth cellular</i>	<i>Bluetooth packet</i>	<i>Bluetooth specification</i>
<i>Bluetooth chips</i>	<i>Bluetooth products</i>	<i>Bluetooth system</i>
<i>Bluetooth connection</i>	<i>Bluetooth qualification</i>	<i>Bluetooth technology(ies)</i>
<i>Bluetooth device(s)</i>	<i>Bluetooth radio(s)</i>	<i>Bluetooth wireless</i>
<i>Bluetooth edition</i>	<i>Bluetooth scatternet(s)</i>	
<i>Bluetooth links</i>	<i>Bluetooth SIG</i>	

4. AGRUPACIONES LÉXICAS O CLUSTERS.

4.1 Distribución de clusters.

Combinaciones	Volumen	Frecuencia máxima
2-cluster	67	41
3-cluster	24	3
4-cluster	17	8

4.2 Inclusión de combinaciones especializadas en unidades más amplias.

<p>Bluetooth packet <i>Bluetooth packet types</i> <i>different Bluetooth packet</i> <i>different Bluetooth packet types</i> Bluetooth radio <i>Bluetooth radio technology</i></p>	<p>Bluetooth qualification <i>Bluetooth qualification program</i> Bluetooth wireless <i>Bluetooth wireless communications</i> <i>Bluetooth wireless technology</i></p>
--	--

4.3 Colocados contiguos de las combinaciones especializadas.

PRE +	COMBINACIÓN	+ POST
<i>fully</i>	<i>integrated Bluetooth</i>	<i>radio, device, FPGA, phone, subsystem, system</i>
	<i>Bluetooth cellular</i>	<i>phone, handsets</i>
<i>attach, contain</i>	Bluetooth chips	
<i>local</i>	<i>Bluetooth connection</i>	<i>organized</i>
<i>integrated, active, each, allow, contemplated, personal, actual</i>	<i>Bluetooth device(s)</i>	<i>establish, per unit, packet, operating</i>
SDK	Bluetooth edition	<i>accelerates, comes</i>
Wi-Fi , <i>lap-tops</i>	<i>Bluetooth-enabled</i>	<i>devices, portable</i>
<i>wireless, technology</i>	<i>Bluetooth links</i>	<i>operate, natively</i>
	<i>Bluetooth mobile</i>	<i>phone</i>
<i>different, types of</i>	<i>Bluetooth packet</i>	<i>types, have</i>
<i>called</i>	<i>Bluetooth products</i>	<i>introductions, receive, Xilinx</i>
	Bluetooth qualification	<i>review, program, body, administrator, test, process</i>
<i>basic, dBm, sources, system, using, typical, future, homeRF, first</i>	<i>Bluetooth radio(s)</i>	<i>components, technology, modules, typically, uses, communications, -on-a-chip</i>
	Bluetooth SOC	<i>digital-to-analog, arbiter</i>
<i>open</i>	<i>Bluetooth specification</i>	<i>enables, defines, Blu99, includes, version, defines</i>
<i>integrated,</i>	<i>Bluetooth system</i>	<i>logic, supports</i>
<i>created, making, medium, way, future</i>	<i>Bluetooth technology(ies)</i>	<i>particularly, focuses, have, does, operates, allows, is optimized facilitates, eliminates, Blu99</i>
<i>contrast, conversely, coolest, non-</i>	<i>Bluetooth wireless</i>	<i>communication(s), link, device, technology(ies), phone, personal, connection</i>

10. Familia de *firewall*. Forma analizada: FIREWALL

1. FRECUENCIA.

KEYWORD	Frec. Telec	Frec. Lacell	Ratio	Término, Chung	KEYNESS	P valor
334. FIREWALL	437	33	50,2930531	ESPECIFICO	1.147,10	0
Miembros relacionados:	F Telec	F Gene	Ratio	Término	Keyness	P valor
FIREWALLS	190	3	240,531993	ESPECIFICO	566,4	0
Familia técnica: Sí, todos los miembros son valorados como específicos.						

2. DISTRIBUCIÓN.

Distribución por áreas	Distribución Keyness	Keykeyword	Nº de textos	Porcentaje
7 (- 2, 5)	3,082	FIREWALL	29	1,75%
<i>Dispersion plot</i>				
Area	Words	Hits	Per 1,000	
082 Esp.Tele	997.727	204	0.20	
3 Telematics	1.205.064	153	0.13	
4 Signal proc.	580.936	35	0.06	
6 Business	373.079	15	0.04	
081 Esp. Sign	867.208	26	0.03	
1 Electronics	722.823	6	0.01	
7 Systems	307.655	1	0.00	

3. COLOCACIÓN.

3.1 Colocados.

N	Collocates	TOTAL	LEFT	RIGHT	L5	L4	L3	L2	L1	*	R1	R2	R3	R4	R5
1	THE	328	210	118	30	21	19	23	117	0	0	30	22	35	31
2	A	203	146	57	13	10	23	22	78	0	0	15	16	15	11
3	AND	114	47	67	15	3	5	11	13	0	31	12	9	8	7
4	TO	96	49	47	12	14	8	15	0	0	9	9	3	12	14
5	OF	68	49	19	7	7	8	18	9	0	1	1	4	6	7
6	IS	64	23	41	5	9	7	2	0	0	19	4	5	8	5
7	FOR	46	27	19	6	5	2	13	1	0	5	5	6	3	0
8	IN	45	25	20	9	4	5	7	0	0	1	6	6	5	2
9	AS	43	23	20	5	4	2	8	4	0	1	4	7	3	5
10	THAT	39	21	18	5	7	4	4	1	0	7	7	1	1	2
11	NETWORK	35	21	14	4	3	3	3	8	0	0	3	3	4	4
12	ARE	34	14	20	6	5	2	0	1	0	2	8	2	5	3
13	SECURITY	34	17	17	6	4	4	3	0	0	5	1	3	6	2
14	CAN	33	11	22	4	1	4	2	0	0	7	7	1	3	4
15	SERVER	32	13	19	3	1	7	0	2	0	12	0	1	3	3
16	BE	29	12	17	3	5	3	1	0	0	0	5	4	3	5
17	WITH	29	14	15	1	3	3	3	4	0	2	2	4	4	3
18	NAT	28	21	7	0	0	2	10	9	0	1	4	0	0	2
19	SERVICES	28	12	16	2	5	3	1	1	0	9	1	2	1	3
20	AN	26	13	13	3	5	4	1	0	0	0	6	4	1	2

21	ON	26	13	13	2	2	0	8	1	0	1	4	1	2	5
22	SOFTWARE	26	9	17	1	4	2	1	1	0	7	1	2	5	2
23	INTERNET	25	10	15	4	3	1	2	0	0	0	0	3	9	3
24	OR	24	10	14	1	0	4	1	4	0	4	3	0	4	3
25	HOST	23	17	6	2	1	2	12	0	0	1	0	0	2	3
26	THIS	23	10	13	7	1	0	0	2	0	0	6	2	2	3
27	VPN	23	9	14	2	1	2	3	1	0	1	6	2	4	1
28	THROUGH	20	16	4	1	0	1	14	0	0	0	3	0	0	1
29	YOU	20	8	12	4	4	0	0	0	0	0	6	5	0	1
30	APPLICATION	19	10	9	2	0	2	6	0	0	0	0	6	2	1
31	BASED	19	17	2	3	0	0	0	14	0	0	1	0	0	1
32	IT	19	9	10	5	3	1	0	0	0	1	3	2	2	2
33	PACKET	19	13	6	2	2	1	8	0	0	0	0	2	1	3
34	ACCESS	18	5	13	3	0	2	0	0	0	0	2	5	0	6
35	IP	18	5	13	0	4	0	0	1	0	0	3	7	3	0

3.2 Colocados significativos.

Nº	Colocados	Z	T	MI
1	GAUNTLET	95,42	2,45	10,57
2	FIREWALL'S	95,42	2,45	10,57
3	IPTABLES	59,23	2,82	8,78
4	STATEFUL	56,81	3,15	8,34
5	NAT	87,07	5,27	8,09
6	FIREWALLS	28,01	3,27	6,20
7	VRF	20,25	2,41	6,14
8	INTRUSION	18,28	2,20	6,10
9	CONSULTING	18,41	2,41	5,87
10	COSINE	24,06	3,26	5,77
11	IPSEC	25,95	3,67	5,65
12	ANTI	15,96	2,39	5,47
13	INSTALLING	14,28	2,18	5,42
14	BEHIND	20,74	3,37	5,24
15	SCRIPT	11,52	2,16	4,83

3.3 Definición del nodo en el diccionario.

1. Definición de *Firewall* en el *Advanced Learner's Dictionary* de Cambridge (2003):

Firewall (noun): (specialized) a device or program that stops people accessing a computer without permission while it is connected to the Internet.

2. Definición de *firewall* en el diccionario especializado Webster:

Firewall (noun):

1. (informal) the application of maximum thrust; "he slammed the throttle to the firewall".
2. Fireproof (or fire-resistant) wall designed to prevent the spread of fire through a building.

Specialty Definition:

The computing world uses the term firewall for a piece of hardware or software put on the network to prevent some communications forbidden by the network policy.

3.4 Colocados contiguos.

a) Colocado + <i>firewall</i> .		
Total colocados: 12	Frecuencia máxima: 14	Frecuencia mínima: 3
- Sustantivos, total 4: <i>network</i> 8, <i>IPtables</i> 4, <i>proxy</i> 4, <i>routers</i> 3.		
- Adjetivos, total 4: <i>corporate</i> 6, <i>personal</i> 6, <i>stateful</i> 5, <i>specific</i> 3.		

- Participios verbales, total 4: <i>based 14, managed 6, filtering 5, routing 3.</i>		
b) Firewall + colocado.		
Total colocados: 14	Frecuencia máxima: 12	Frecuencia mínima: 3
- Sustantivos, total 13: <i>server 12, rules 9/rule 8, services 9, system 9/systems 6, software 7, performance 6, protection 6, functionality 4, policie(s) 4/policy 3, technology 3.</i>		
- Siglas, total 1: <i>S 5.</i>		

3.5 Identificación de colocados contiguos especializados en las líneas de concordancia.

*of corresponding tuples. In addition to the prefixes, each **firewall rule** has an action, which is to accept or deny. An Script" "Installing the Firewall Rule Set" "Removing the **Firewall Rule Set**" "Verifying Firewall Services" "Firewall "Creating a Host-Based Firewall Script" "Installing the **Firewall Rule Set**" "Removing the Firewall Rule Set" "Verifying not use a fixed port number for the data connection. The **firewall rule** for the FTP case creates a new entry in the stateful disk, a slow operation that causes message loss when the **firewall rule** set generates a lot of messages. The rule set design of resulting policy tries would mirror those of real-world **firewall rule** sets. Policy sizes ranged from 50 rules to 500 rules, a tuple, a path from the root node to a leaf represents a **firewall rule**, as seen in figure 1. To create a policy trie T, rules (list) of n rules, denoted as R = {r1, r2, ..., rn}. Similar to a **firewall rule**, a packet (IP datagram) d can be viewed as an to change filter settings in the kernel are privileged. Add **firewall rules** one rule at a time, using multiple iptables command be used to enhance existing security measures such as **firewall rules** and IDS configurations. The Evolution of Network organized as follows: Section 2 describes the models for **firewall rules** and a standard (list-based) security policy. The policy has been traditionally defined as an ordered list of **firewall rules** [23], as seen in table 1. A rule r can be viewed rule chain named discard to process all packets that our **firewall rules** do not pass. Depending on your environment, Figure 3. Push-down policy trie representation of the **firewall rules** given in table 1. Simulations were conducted using can access be granted to the internal network. The **firewall rules** may also be used to limit where traffic originating modules. Prepare the Firewall. First remove the previous **firewall rules**, then delete all user-defined chains. These tasks a good idea to log every single packet that is rejected by **firewall rules**. Most IP networks tend to be noisy environments,*

Concordancias de *firewall rule(s)*.

a) Colocado + *firewall*:

<i>based firewall</i>	<i>managed firewall</i>	<i>specific firewall</i>
<i>corporate firewall</i>	<i>network firewall</i>	<i>stateful</i> <i>firewall</i>
<i>filtering firewall</i>	<i>personal firewall</i>	
<i>IPtables</i> <i>firewall</i>	<i>proxy firewall</i>	

b) *Firewall* + colocado:

<i>firewall functionality</i>	<i>firewall rule(s)</i>	<i>firewall system(s)</i>
<i>firewall performance</i>	<i>firewall server</i>	<i>firewall technology</i>
<i>firewall policy(ies)</i>	<i>firewall services</i>	
<i>firewall protection</i>	<i>firewall software</i>	

4. AGRUPACIONES LÉXICAS O CLUSTERS.

4.1 Distribución de clusters.

Combinaciones	Volumen	Frecuencia máxima
2-cluster	33	6
3-cluster	53	12
4-cluster	9	5

4.2 Inclusión de combinaciones especializadas en unidades más amplias.

<p><i>based firewall</i> <i>based firewall script</i> <i>network firewall</i> <i>network firewall performance</i> <i>proxy firewall</i> <i>application proxy firewall</i> <i>gauntlet application proxy firewall</i></p>	<p><i>firewall rule</i> <i>firewall rule set</i> <i>firewall services</i> <i>verifying firewall service</i></p>
--	--

4.3 Colocados contiguos de las combinaciones especializadas.

PRE +	COMBINACIÓN	+ POST
<i>Host-</i>	<i>based firewall</i>	<i>functionality, script, setup, similarly, solution, systems</i>
	<i>corporate firewall</i>	<i>features, system</i>
<i>packet-</i>	<i>filtering firewall</i>	
	<i>IPtables</i> <i>firewall</i>	<i>maintains, uses</i>
<i>centrally, offer</i>	<i>managed firewall</i>	<i>services</i>
<i>increase, improving</i>	<i>network firewall</i>	<i>performance, technologies</i>
<i>software's, Norton, Pro</i>	<i>personal firewall</i>	<i>software</i>
<i>application</i>	<i>proxy firewall</i>	
<i>extranet</i>	<i>specific firewall</i>	<i>-related</i>
<i>CoSine</i> - <i>originated</i>	<i>stateful</i> <i>firewall</i>	
<i>host-based</i>	<i>firewall functionality</i>	
<i>improving, increase, network</i>	<i>firewall performance</i>	
<i>adjust, different, model, virtualized, rule</i>	<i>firewall policy(ies)</i>	<i>NAT</i> , <i>enforced</i>
<i>extended, extra, layer 2</i>	<i>firewall protection</i>	<i>capabilities</i>
<i>real-world, add, previous</i>	<i>firewall rule(s)</i>	<i>set(s), given</i>
<i>deployed</i>	<i>firewall server(s)</i>	<i>represents</i>
<i>advanced, managed, verifying</i>	<i>firewall services</i>	
<i>embedded, personal</i>	<i>firewall software</i>	<i>service, application</i>
<i>corporate, few, perimeter-based</i>	<i>firewall system(s)</i>	<i>operating, placed, including, comprise</i>
<i>network, sophisticated, today's</i>	<i>firewall technology(ies)</i>	

11. Familia de *crossstalk*. Forma analizada: CROSSTALK

1. FRECUENCIA.

KEYWORD	Frec. Telec	Frec. Lacell	Ratio	Término, Chung	KEYNESS	P valor
943. CROSSTALK	90	0	inf/esp	inf/esp	282,3	0
Término especializado.						

2. DISTRIBUCIÓN.

Distribución por áreas	Distribución Keyness	Keykeyword	Nº de textos	Porcentaje
6 (- 5, 6, 7)	081	CROSSTALK	10	0,6%
<i>Dispersion plot</i>				
Area	Words	Hits	Per 1,000	
081 Esp. Sign	867.208	40	0.05	
4 Signal proc.	580.936	25	0.04	
2 Ar. Comp	329.643	8	0.02	
1 Electronics	722.823	10	0.01	
082 Esp.Tele	997.727	7	0.01	
3 Telematics	1.205.064	2	0.00	

3. COLOCACIÓN.

3.1 Colocados.

N	Collocates	TOTAL	LEFT	RIGHT	L5	L4	L3	L2	L1	*	R1	R2	R3	R4	R5
1	THE	51	28	23	6	5	7	2	8	0	0	6	6	2	9
2	AND	47	24	23	5	4	1	4	10	0	8	8	2	3	2
3	OF	30	26	4	2	3	6	7	8	0	0	0	1	2	1
4	IN	22	4	18	0	2	1	1	0	0	8	6	1	2	1
5	IS	20	5	15	0	2	2	1	0	0	4	3	3	0	5
6	TO	19	9	10	1	1	3	2	2	0	0	3	3	1	3
7	OPTICAL	17	13	4	2	0	0	1	10	0	0	1	0	1	2
8	A	12	5	7	1	0	2	0	2	0	0	1	3	2	1
9	NOISE	11	4	7	0	0	1	1	2	0	6	0	1	0	0
10	END	9	7	2	0	1	0	0	6	0	0	0	1	0	1
11	ARE	8	3	5	1	1	0	1	0	0	1	0	1	1	2
12	BETWEEN	8	0	8	0	0	0	0	0	0	4	1	0	0	3
13	FOR	8	5	3	1	0	0	2	2	0	0	1	0	2	0
14	ON	8	3	5	1	1	1	0	0	0	1	1	2	0	1
15	THIS	8	4	4	2	2	0	0	0	0	0	3	0	1	0
16	FROM	7	5	2	0	1	1	1	2	0	1	0	1	0	0
17	NEAR	7	6	1	1	0	0	5	0	0	0	0	0	1	0
18	THAT	6	3	3	3	0	0	0	0	0	1	0	0	1	1
19	UP	6	3	3	2	0	0	0	1	0	0	0	0	2	1
20	AT	5	1	4	0	1	0	0	0	0	2	0	0	1	1
21	DUE	5	3	2	0	2	1	0	0	0	1	0	0	1	0
22	HAVE	5	3	2	0	1	1	0	1	0	0	1	1	0	0
23	MINIMIZING	5	4	1	2	0	2	0	0	0	0	1	0	0	0
24	NETWORKS	5	0	5	0	0	0	0	0	0	0	0	1	1	3
25	POWER	5	3	2	0	1	0	2	0	0	1	0	0	0	1

26	RADIO	5	0	5	0	0	0	0	0	0	0	0	2	3	0
27	REFERENCE	5	5	0	1	0	1	1	2	0	0	0	0	0	0
28	SIGNAL	5	3	2	0	0	2	0	1	0	0	0	1	1	0

3.2 Colocados significativos.

* Estos son los colocados más significativos, pero no cumplen con todos los requisitos.

Nº	Colocados	Z	T	MI
1	MINIMIZING	36,27	2,23	8,05
	NEAR	21,51	2,61	6,09
3	OPTICAL	24,41	4,01	5,21
4	NOISE	19,27	3,22	5,16
5	REFERENCE	9,48	2,12	4,32
6	RADIO	8,78	2,11	4,12
7	END	10,55	2,79	3,84
8	DUE	7,56	2,07	3,74
9	POWER	4,62	1,87	2,61
10	NETWORKS	4,16	1,81	2,40
11	SIGNAL	3,61	1,73	2,13

3.3 Definición del nodo en el diccionario.

1. El *Advanced Learner's Dictionary* de Cambridge (2003) no registra la definición de *crosstalk*.
2. Definición de *crosstalk* en el diccionario especializado Webster:

Crosstalk (noun):

1. *The presence of an unwanted signal via an accidental coupling.*

Specialty Definition:

Crosstalk: *Interference caused by two signals becoming partially superimposed on each other due to electromagnetic (inductive) or electrostatic (capacitive) coupling between the conductors carrying the signals. A common example of crosstalk is where the magnetic field from changing current flow in one wire induces current in another wire running parallel to the other, as in a transformer.*

3.4 Colocados contiguos.

a) Colocado + <i>crosstalk</i> .		
Total colocados seleccionados: 4	Frecuencia máxima: 2	Frecuencia mínima: 3
- Sustantivos, total 3: <i>end 6, noise 2, reference 2</i> .		
- Adjetivos, total 1: <i>optical 10</i> .		
b) <i>Crosstalk</i> + colocado.		
Total colocados seleccionados: 1	Frecuencia máxima: 6	Frecuencia mínima: 6
- Sustantivos, total 36: <i>noise 6</i> .		

3.5 Identificación de colocados contiguos especializados en las líneas de concordancia.

tive Coupling. Tetsuhisa Mido, Kunihiro Asada. Crosstalk noise in long interconnections is studied in terms of (i) routing on outside layers, (ii) routing on inside layers, (iii) routing on outside layers, and (iv) routing on inside layers. 3D.3 Crosstalk Noise in High Density and High Speed I/O Tools. We will also present some techniques for crosstalk noise avoidance and delay matching that optimization is essential for this low-power design. Crosstalk noise and delay optimizations are unavailing to detect dynamic errors due to reflection and crosstalk noise in combinational circuits is presentation of Dynamic Errors due to Reflection - and Crosstalk-Noise. J. Schrage. A new algorithm for

Concordancias de *crosstalk noise*.

a) Colocado + *crosstalk*:
end crosstalk
optical crosstalk
reference crosstalk

b) *Crosstalk* + colocado:
crosstalk levels
crosstalk noise
crosstalk problems
crosstalk reduction

4. AGRUPACIONES LÉXICAS O CLUSTERS.

4.1 Distribución de clusters.

Combinaciones	Volumen	Frecuencia máxima
2-cluster	14	11
3-cluster	6	4
4-cluster	3	3

4.2 Inclusión de combinaciones especializadas en unidades más amplias.

<i>optical crosstalk</i> <i>impact of optical crosstalk</i> <i>optical crosstalk in WDM</i>

4.3 Colocados contiguos de las combinaciones especializadas.

PRE +	COMBINACIÓN	+ POST
<i>near-, far</i>	<i>end crosstalk</i>	<i>NEXT, FEXT</i>
<i>effect of, impact of, due to, out-of-band</i>	<i>optical crosstalk</i>	<i>levels, are obtained, in WDM</i>
<i>Input, challenge of</i>	<i>reference crosstalk</i>	
<i>component, optical</i>	<i>crosstalk levels</i>	
<i>techniques for</i>	<i>crosstalk noise</i>	<i>avoidance</i>
<i>potential</i>	<i>crosstalk problems</i>	<i>resulting</i>
<i>consideration of</i>	<i>crosstalk reduction</i>	

10.4 Análisis detallado de formas académicas.

En el apartado 9.9.3 se hizo referencia al estudio de Wang y Nation (2004) con la intención de ratificar la validez y fiabilidad del listado de vocabulario académico de Coxhead (2000). Dicho estudio fundamenta la decisión adoptada de aceptar que las palabras clave identificadas en el corpus, no estarán provistas de significados adicionales siempre que coincidan con las formas registradas en el listado académico. Sin embargo, también se plantean ciertas dudas, surgidas a lo largo de los diversos análisis realizados, sobre una posible adaptación del significado de las formas académicas *network*, *router*, *data* y *layer*, dada la estrecha relación que parecen mantener con el área de las telecomunicaciones. Asimismo, las familias a las que pertenecen estas formas se componen en su mayoría de miembros no etiquetados como académicos y valorados como específicos. El conjunto de estos factores suscita realizar el análisis detallado correspondiente con el fin de resolver las dudas planteadas. Por lo tanto, a continuación se presentan las fichas correspondientes al análisis de las formas académicas, siguiendo el modelo establecido con *bandwidth*. En el Apéndice VII, se hallan las fichas de todas las formas analizadas, junto a los listados completos de los colocados, colocados significativos, líneas de concordancia y clusters. Se incluye en este mismo apéndice la ficha de *data*.








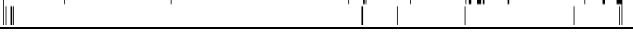

1. Familia de *network*. Forma analizada: NETWORK

1. FRECUENCIA

KEYWORD	Frec. Telec	Frec. Lacell	Ratio	Término, Chung	KEYNESS	P valor
1. NETWORK / Académica	16.649	1.686	37,5034384	NO	41.784,60	0
Miembros relacionados:	F Telec	F Gene	Ratio	Término	Keyness	P valor
NETWORKS / Académica	5.832	463	47,8384421	NO	15.204,90	0
NETWORKING / Académica	1.030	94	41,6149977	NO	2.628,20	0
INTERNETWORK	195	1	740,585347	ESPECIFICO	599,5	0
SUBNETWORK	133	0	inf/esp	inf/esp	417,1	0
NETWORKED / Académica	141	27	19,8333398	NO	306,7	0
INTERNETWORKING	82	3	103,808544	ESPECIFICO	232,6	0
INTERNETWORKS	53	0	inf/esp	inf/esp	166,2	0
SUBNETWORKS	45	0	inf/esp	inf/esp	141,1	0
MULTINETWORK	12	0	inf/esp	inf/esp	37,6	0
FOUNDRYNETWORKS	7	0	inf/esp	inf/esp	22	0,000003
ADNETWORK	6	0	inf/esp	inf/esp	18,8	0,000014
NETWORKMESSENGER	4	0	inf/esp	inf/esp	12,5	0,000397

Familia técnica: No, estadística y formalmente.

2. DISTRIBUCIÓN

Distribución por áreas	Distribución Keyness	Keykeyword	Nº de textos	Porcentaje
9	3, 082	NETWORK	594	35,91%
<i>Dispersion plot</i>				
Area	Words	Hits	Per 1,000	
082 Esp.Tele	997.683	6.667	6.68	
3 Telematics	1.204.955	4.334	3.60	
081 Esp.Sign	867.175	2.762	3.18	
4 Signal proc.	580.890	1.012	1.74	
1 Electronics	722.778	628	0.86	
6 Business	373.043	380	1.24	
7 Systems	307.662	72	0.23	
2 Ar. Comp	329.605	65	0.50	
5 Materials	101.232	10	0.10	

3. COLOCACIÓN

3.1 Colocados.

N	Collocates	TOTAL	LEFT	RIGHT	L5	L4	L3	L2	L1	*	R1	R2	R3	R4	R5
1	THE	12042	7847	4195	1098	935	881	1674	3259	0	24	1166	1055	959	991
2	A	5245	3456	1789	420	475	683	849	1029	0	31	447	480	407	424
3	TO	4251	2411	1840	518	471	621	609	192	0	262	400	338	402	438
4	AND	4209	1884	2325	453	372	384	347	328	0	417	813	354	368	373
5	OF	3987	2642	1345	405	443	529	829	436	0	159	143	205	384	454
6	IN	2867	1596	1271	241	301	427	540	87	0	152	376	252	237	254
7	IS	2446	786	1660	275	249	139	105	18	0	423	319	321	313	284
8	FOR	1854	1081	773	230	198	192	267	194	0	100	223	152	134	164

9	ON	1357	978	379	124	121	254	414	65	0	36	84	72	87	100
10	THAT	1298	573	725	142	122	109	133	67	0	143	173	137	122	150
11	AS	1143	488	655	123	115	102	104	44	0	92	170	153	130	110
12	BY	1008	594	414	71	70	275	135	43	0	59	117	89	73	76
13	WITH	1008	435	573	94	81	101	112	47	0	146	113	98	107	109
14	ARE	1002	356	646	138	125	54	26	13	0	45	157	148	176	120
15	AN	993	633	360	98	104	145	286	0	0	2	84	94	86	94
16	THIS	839	285	554	60	68	52	33	72	0	8	221	122	91	112
17	OR	827	424	403	82	84	95	92	71	0	84	113	61	74	71
18	BE	805	300	505	126	125	33	16	0	0	4	123	157	121	100
19	CAN	718	193	525	71	56	53	12	1	0	111	115	108	92	99
20	MANAGEMENT	690	183	507	39	40	38	22	44	0	352	41	49	32	33
21	FROM	619	352	267	71	62	80	115	24	0	54	59	39	48	67
22	IT	607	187	420	64	61	44	12	6	0	12	149	104	90	65
23	ACCESS	604	341	263	56	52	77	56	100	0	122	21	37	44	39
24	LAYER	592	143	449	30	40	38	28	7	0	276	86	14	41	32
25	BASED	520	192	328	35	28	37	31	61	0	178	36	29	39	46
26	SERVICES	483	168	315	41	41	27	50	9	0	121	43	72	43	36
27	DATA	476	271	205	60	53	58	34	66	0	12	33	53	57	50
28	IP	473	322	151	57	38	28	31	168	0	9	33	33	36	40
29	YOUR	440	365	75	19	14	27	67	238	0	0	13	25	21	16
30	TRAFFIC	439	183	256	39	74	56	11	3	0	132	18	38	38	30
31	AT	435	203	232	56	37	24	67	19	0	31	68	42	49	42
32	WHICH	428	147	281	44	31	27	27	18	0	20	117	56	51	37
33	WIRELESS	414	297	117	30	25	30	20	192	0	1	13	40	35	28
34	SERVICE	403	185	218	42	60	33	41	9	0	65	31	53	39	30
35	HAS	393	100	293	47	23	18	12	0	0	87	68	29	49	60

3.2 Colocados significativos.

Nº	Colocados	MI	Z	T
1	VISITED	15,28161256	4,9114	3,2751
2	VINES	14,4163144	4,1088	3,6218
3	UNPLUGGED	22,39851168	4,4053	4,6922
4	TRANSPARENTLY	16,74603848	4,848	3,5767
5	TOPOLOGY	37,17128537	12,73	3,0916
6	TELECITIES	18,76956979	4,6042	4,0547
7	SURVIVABILITY	11,96345632	3,6314	3,44
8	SUBNET	26,07527825	9,0876	3,0413
9	RECURRENT	15,99647589	4,0759	3,9451
10	PSTN	18,89955175	5,9151	3,3517
11	PROTECTORS	26,80571818	4,7454	4,9958
12	PROTECTING	14,68206729	4,74	3,2622
13	PERIMETER	13,40868371	4,377	3,2303
14	NONEXTENDED	23,72377919	4,115	5,0547
15	NGSO	18,04574273	4,32	4,1251
16	NEWSLETTER	44,42558722	10,002	4,3023
17	NEURO	30,2462992	6,4718	4,449
18	NEURAL	39,92121239	10,419	3,8759
19	NETWORK'S	57,37242168	9,3376	5,2385
20	NBMA	20,71961953	5,9469	3,6016
21	GPRS	26,47181462	9,3141	3,0138
22	FEEDERS	16,45706125	4,0897	4,0173
23	CIVIC	12,0741995	3,9033	3,2583
24	CITYLINK	11,45056268	3,3051	3,5853

25	CARRIER'S	11,17119863	3,3959	3,4358
26	APPN	14,66309486	4,6692	3,3018
27	APPLIANCE	31,16965716	7,427	4,1386
28	ADMINISTRATORS	24,73393468	8,0473	3,2398
29	ADMINISTRATOR	24,49604579	8,4862	3,0586
30	AD	25,23067577	8,7524	3,0547

3.3 Definición del nodo en el diccionario.

- Definición de *network* en el *Advanced Learner's Dictionary* de Cambridge (2003):

Network (n): *a large system consisting of many similar parts that are connected together to allow movement or communication between or along the parts or between the parts and a control centre: A television network; a road/rail network; a computer network.*

Network (T.,v): *to connect computers together so that they can share information: Our computer system consists of about twenty personal computers networked to a powerful file-server.*

Network (I. v): *to meet people who might be useful to know, especially in your job: I don't really enjoy these conferences, but they're a good opportunity to do some networking.*

- Definición de *network* en el diccionario especializado Webster:

Network (noun):

- An intricately connected system of things or people; "a network of spies"; "a web of intrigue".*
- Communication system consisting of a group of broadcasting stations that all transmit the same program simultaneously.*
- An interconnected or intersecting configuration or system of components; "there was a hole in the network where some of the strands were broken"; "he used a copper frame with copper meshing".*

Network (noun): *Communicate with and within a group.*

Specialty Definition:

In computing:

Network Hardware and software data communication systems. The OSI seven layer model attempts to provide a way of partitioning any computer network into independent modules from the lowest (physical) layer to the highest (application) layer. Many different specifications exist at each of these layers. Networks are often also classified according to their geographical extent: local area network (LAN), metropolitan area network (MAN), wide area network (WAN) and also according to the protocols used.

In electrical engineering:

A grouping of lines and of other electrical equipment connected for the purpose of conveying electricity from generating stations to the ultimate consumer.

In post & telecom:

- A group of broadcast stations connected by cable or microwave relay for simultaneous broadcasting of a program.*
- One or several interconnected stages forming part of the overall switching network of an automatic switching system and performing one specific function.*

3.4 Colocados contiguos.

a) Colocado + <i>network</i> .		
Total colocados seleccionados: 84	Frecuencia máxima: 195	Frecuencia mínima: 15
<ul style="list-style-type: none"> - Sustantivos, total 27: <i>home 104, telephone 84, ring 78, core 69, relay 67, backbone 66, data 66, enterprise 56, Ethernet 38, destination 39, communications 33/communication 27, transit 31, broadcast 30, carrier 26, transport 25, distribution 24, star 23, interconnection 21, computer 18, telecityies 17, radio 17, mesh 16, Appletalk 15, broadband 15, generation 15, sensor 15.</i> - Adjetivos, total 33: <i>wireless 192, optical 153, neural 113, mobile 106, private 84, hoc 75, same 67, cellular 63, physical 61, corporate 61, different 54, entire 53, digital 44, fuzzy 44, internal 43, single 40, intelligent</i> 		

<i>AppleTalk network</i>	<i>mesh network</i>	<i>satellite network</i>
<i>ATM network</i>	<i>mobile network</i>	<i>sensor network</i>
<i>backbone network</i>	<i>modular network</i>	<i>SP network</i>
<i>broadband network</i>	<i>MPLS network</i>	<i>star network</i>
<i>broadcast network</i>	<i>multiaccess network</i>	<i>star network</i>
carrier network	<i>multifiber network</i>	<i>stub network</i>
<i>cellular network</i>	<i>NBMA network</i>	<i>switched network</i>
civic network	neural network	<i>TCP network</i>
<i>communication network</i>	<i>NGSO network</i>	<i>telecommunications network</i>
<i>computer network</i>	<i>optical network</i>	telecities network
<i>core network</i>	<i>OSI network</i>	<i>telephone network</i>
<i>corporate network</i>	<i>P2P-network</i>	<i>terrestrial network</i>
<i>data network</i>	<i>packet network</i>	<i>transit network</i>
<i>digital network</i>	<i>peer-to-peer network</i>	<i>transport network</i>
<i>distribution network</i>	<i>phone network</i>	<i>UMTS network</i>
<i>Ethernet network</i>	<i>physical network</i>	<i>voice network</i>
<i>fuzzy network</i>	<i>point-to-point network</i>	<i>WDM network</i>
<i>global network</i>	<i>radio network</i>	<i>wired network</i>
GPRS network	<i>relay network</i>	<i>wireless network</i>
<i>GSM network</i>	<i>ring network</i>	<i>WLAN network</i>
<i>IP network</i>	<i>RIP network</i>	

b) Network + colocado:

<i>network access</i>	<i>network equipment</i>	<i>network operator</i>
<i>network adapter</i>	<i>network failure</i>	<i>network performance</i>
<i>network address</i>	<i>network foundation</i>	<i>network processor</i>
<i>network administrator</i>	<i>network infrastructure</i>	<i>network protectors</i>
<i>network architecture</i>	<i>network installation</i>	<i>network provider</i>
<i>network availability</i>	<i>network interface</i>	<i>network resources</i>
<i>network bandwidth</i>	<i>network latency</i>	<i>network security</i>
<i>network backbones</i>	<i>network layer</i>	<i>network segment</i>
<i>network capacity</i>	<i>network level</i>	<i>network service</i>
<i>network component</i>	<i>network links</i>	<i>network switches</i>
<i>network configuration</i>	<i>network load</i>	<i>network systems</i>
<i>network congestion</i>	<i>network LSA/LSAs</i>	<i>network technologies</i>
<i>network connection</i>	<i>network management</i>	<i>network throughput</i>
<i>network controllers</i>	<i>network mask</i>	<i>network topology</i>
<i>network design</i>	<i>network model</i>	<i>network traffic</i>
<i>network device</i>	<i>network nodes</i>	
<i>network environment</i>	<i>network operation</i>	

4. AGRUPACIONES LÉXICAS O CLUSTERS.

4.1 Distribución de clusters.

Combinaciones	Volumen	Frecuencia máxima
2-cluster	690	473
3-cluster	1368	471
4-cluster	1431	41
5-cluster	530	7
6-cluster	296	30

4.2 Inclusión de combinaciones especializadas en unidades más amplias.

<p><i>network access</i> <i>network access layer</i> <i>network access control</i> <i>network access method</i> <i>network access point(s)</i> <i>network access server</i> <i>network access server NAS</i> <i>level of network access</i></p>	<p>MPLS network <i>single MPLS network</i> <i>IP MPLS network</i> NBMA network <i>broadcast or NBMA network</i> neural network <i>neural network model(s)</i> <i>artificial neural network</i></p>
--	--

<p>network address network address translation availability network address network address translation NAT</p> <p>network architecture system network architecture network architecture SNA digital network architecture open- network architecture ONA system network architecture SNA multiwavelength network architecture layered multiwavelength network architecture layered multiwavelength network architecture point</p> <p>network connection physical network connection</p> <p>network device ATM network device(s) programmable network device(s)</p> <p>network environment simulated network environment</p> <p>network equipment optical network equipment</p> <p>network layer network layer protocol(s) network layer address(es) OSI network layer multiple network layer network layer multicast destination network layer network layer header network layer routing network layer creates network layer functions network layer tree primary network layer multiple network layer protocols primary network layer routing network layer routing protocol network layer protocol configuration network layer multicast addresses destination network layer address configuring different network layer underlying network layer protocols network layer protocol type network layer protocol configuration negotiation primary network layer routing protocol</p> <p>network operation network operation center network operation centers NOCs</p> <p>network operators mobile network operators wireless network operators providers and network operators generation wireless network operators network operators and service providers third generation wireless network operators</p> <p>network service network service providers intelligent network service data network services network service access network service division network service security managed data network service network service access point transaction network service division network service security layer</p>	<p>spline neural network neural network-based</p> <p>OSI network OSI network layer OSI network management</p> <p>relay network frame relay network simple frame relay network</p> <p>ring network token ring network MCC ring network IBM token ring network token ring network and IEEE</p> <p>telephone network telephone network PSTN switched telephone network public telephone network public switched telephone network PSTN public switched telephone network telephone network and the internet transport network optical transport network</p> <p>WDM network optical WDM network</p> <p>wireless network wireless network operators available wireless network safety wireless network public safety wireless network highly available wireless network wireless network PSWN program wireless network interface card third generation wireless network safety wireless network PSWN multimedia and wireless network generation wireless network operators public safety wireless network PSWN public safety wireless network program third generation wireless network operators second and third generation wireless network</p> <p>network infrastructure mobile network infrastructure network infrastructure optimized</p> <p>network installation network installation procedures server network installation secure server network installation access manager network installation WLAN secure server network installation</p> <p>network interface network interface card(s) network interface UNI network interface PNNI network-network interface network interface layer user- network interface network interface agreements network interface LAN network interface NNI network interface cards NICS network-to- network interface private network-network interface wireless network interface card user to network interface the network interface unit subscriber network interface SNI network interface unit NIU</p>
---	--

<p>highly available network services network service access point NSAP network system access network system optical network systems WLAN access network system Xerox network system XNS network topology network topology changes underlying network topology network topology knowledge network topology map picture of network topology network model neural network model fuzzy network model TCP/IP network model network node active network node network node NN network security network security policy network security policy management</p>	<p>network management network management system network management protocol simple network management network management software IBM network management network management layer customer network management network management architecture network management functions network management tools OSI network management RAD's network management portfolio simple network management protocol network management system NMS network management and service assurance security for network management network management functional areas customer network management CNM simple network management protocol SNMP network segment network segment basis physical network segment associated network segment a per network segment basis</p>
--	---

4.3 Colocados contiguos de las combinaciones especializadas.

pre +	COMBINACIÓN	+ post
single, carrier , public, private, common, general, core, non-, underlying, well-dimensioned, prototypical, end, mode	ATM network	based, behavior, carrying, comprises, consists, devices(s), infrastructure, interconnections, interface(s), management, look(s), performance, participating, topology
Alcatel's, ad hoc, wired-line, ABR, FR, modelled, WDM, core, OPS, routed, IP, packet, switched, ATM, inter-metro, simulate	backbone network	contains, complemented, dominating, generally, ISPs, model, offering, reliable, serves, requiring, using, via
terrestrial, multi-hop, digital, micro-, existing, comparable, global, satellite-based, virtual, area, traditional, multi-layered	cellular network	ability, architecture, arises, BS, comprising, converts, using, environment, fraud, VCN, incorporated, infrastructure, involves, operators, systems
IP, internet-, GPRS , MPLS, IP-based, unified, integrated, loaded, resulting, central, Ericsson's, ensuring, valuable, ATM, high-traffic	core network	architecture, CN, demarcation, equipment, evolving, foundation, including, operator(s), platform, products, provides, relays, reliability, resources, switch
switched, fast, metro, gigabit, industrial, users', corporate, shared	Ethernet network	adapters, based, becomes, cards, carries, control, elements, integrated, interface, running, topology(ies) , users, devices
service(s), OSI, carrier's , proprietary	digital network	architecture, B-ISDN, interfaces, market, basic, rates, services
wireless, traditional, WAN, typical, packet-switched, private, CM-5, private, existing, services, packet, public, cellular, switched, managed, high-speed, core, large	data network	background, elements, expands, gear, identification, like, infrastructure, introduces, management, needs, offering, protocols, providers, PSDN, PSTN , services, switches, topologies , usage
neuro- ,	fuzzy network	-based, controller, model(s), modelling, process, representations
core, corresponding, per, radio, Vodafone, existing, new, local,	GSM network	broadcast, capacity, components, elements, operator(s), OTA

<i>generic, Srbija's, first, generation, telephony</i>		
<i>premium, tree, destination, all, shared, WorldCom's, broadband, providing, logical, unloaded, company's, existing, CityLink, satellite, private, site, based, all-, future, voice-over-, packets, typical, same, single, assigned, multiple, registered, larger, IPNOS, services, cellular, distributed, global</i>	<i>IP network</i>	<i>access, accounting, address(es), connectivity, consistent, deployment, designer, devices, different, increased, infrastructure, M3I, management, mask, messages, number(s), operating, packets, protocol, subnet</i>
<i>cellular, home, land, emerging, visited, CDMA, 3G, underlying, driving, 2.5G, own, secure, MNO, 3GSM, sign up, partnership, available, private, land, based, traditional, foreign, area, handover, commercial</i>	<i>mobile network</i>	<i>architecture(s), carrier(s), community, deployments, gain, security, enhanced, environments, evolution, infrastructure, keeping, market, members, scenario, nationally, operator(s), committed, signed, provide, view, PLMN, roaming,</i>
<i>generic, TMN, MIBs, dedicated</i>	<i>OSI network</i>	<i>address, layer, management, model, protocol, service, using</i>
<i>frame</i>	<i>relay network</i>	<i>appears, connections, connects, consists, edge, equipment, experiences, implementation, provider, services, switches</i>
<i>interconnected, multi-MCC, WDM, single-MCC, token, Cambridge, fiber-optic, fixed-conversion, bidirectional, unidirectional, local, multifiber, passing</i>	<i>ring network</i>	<i>communication, badly, enables, requires, specifies, stations, topology, components</i>
<i>public, switched, mobile, cellular, regular, local, giant, India's, global</i>	<i>telephone network</i>	<i>T-1, connection, data, during, PSU, infrastructure(s), needs, signalling, overload, PSTN</i>
<i>optical, WDM-based, high-bandwidth, MPLS, IBM, metro, 3G, multiwavelength, regional, multi-layer, voice</i>	<i>Transport network</i>	<i>characteristics, consists, design, facilities, MWTN, OTN</i>
<i>study, all-optical, optical, SONET-over-, hierarchical, state-of-the-art, access</i>	<i>WDM network</i>	<i>architectures, constructions, design, equipment, standards</i>
<i>offer, good, available, external, CPE, proper, often, safety, Nortel, Cisco's, integrated, high-speed, information, install, local, area, half-duplex, RF, 3G, high-capacity, generation, AirPort, DSSS, editing, configure, intelligent, Cingular's, airport's, default, enables, short-range, global, entire, reduce, multi-hop,</i>	<i>wireless network</i>	<i>access, applications, barrier, architecture, card(s), clients, communications, connection(s), design(s), designers, devices, equipment, features, group, interface, installations, key, makes, means, meets, module, needs, node, operation, option, operator(s), PC, performance, policy(ies), PSWN, round, protection, routers, security, settings, SM, standard, technologies, throughput, topology, viewpoint</i>
<i>TCP/IP, area, rights-based, Bell, computer's, configure, control, lower-cost, CSMA, current, on-demand distribution, existing, fixed, guarantee, internet, IP, key, optional, token-passing, provide, requests, require, school, seamless, high-speed</i>	<i>network access</i>	<i>server, layer, method, points, associated, overhead, required, switches, control, products, node, service, across, technologies, layer, link, device, system, security, software</i>
<i>AppleTalk, high-availability, called, destination, entire, IP, IPX, loopback, LSA's, offer, OSI, own, private, rulebase, source, permanent, specified</i>	<i>network address</i>	<i>translation, format, consists, port, families, space, assignment, skips, numbers, refers, field</i>

<p><i>access, cellular, daisy-chain, considerable, control, core, digital efficient, flat, GSM/GPRS, HFC, ad-hoc, interconnecting, layered, LMDS, multiwavelength, open-, optical, passing, peer-to-peer, systems, redundant, robust, UMTS, ADSL, mobile, wireless</i></p>	<p><i>network architecture</i></p>	<p><i>reduces, data, DNA, comprises, UCAN, consists, point, ONA, design, traffic, SNA, routing, protocols</i></p>
<p><i>emulated, fixed, multiple, SDN, new, online, own, permanent, physical, private, single, T1, telephone, wireless, regular, fast, party, dedicated</i></p>	<p><i>network connection(s)</i></p>	<p><i>capability, profiles, using, requirements, point, interface, sharing, link</i></p>
<p><i>ATM, non-ATM, caused, net, remote, tedious, track, Cisco, physical, connect(s), cooperating, enables, help(s), intermediate, IP, modeling, optical, organize, remaining, end-user, WAN</i></p>	<p><i>network device(s)</i></p>	<p><i>implies, vendors, technology, software, consolidation, programming, interface, performance, accessing, increases, known, contend, generate, locate, hosts, run, supporting</i></p>
<p><i>control, existing, gigabit, global, physical, shared-, simulated, studied, fault-tolerant, changing, realistic, experimental, home</i></p>	<p><i>network environment</i></p>	<p><i>uses, consisted</i></p>
<p><i>advanced, assist, filter-based, certain, core, customers, future, next-generation, iconrol, installing, mesh, optical, relay, routing, advanced, housing, WDM, submitted, UMTS, wireless</i></p>	<p><i>network equipment</i></p>	<p><i>integration, futures, makers, manufacturers, specifications, offerings, operating, designs, vendors, suppliers, platforms, MIBs, providers, evaluation, needed, side, segment</i></p>
<p><i>additional, advanced, AT&T's, ATM, broadband, building, cellular, Cisco, communications, data, deliver, end-to-end, Wi-Fi, installed, mobile, need, optic, primary, reduce, high-speed, voice, wireless, transparent, computer, packet, telephone</i></p>	<p><i>network infrastructure(s)</i></p>	<p><i>consisting, solutions, helps, consolidation, replaced, cripples, switches, deployments, developments, compliance, strategy, capable, transparently</i></p>
<p><i>manager, perform, server, manager, existing, wireless</i></p>	<p><i>network installation(s)</i></p>	<p><i>procedures</i></p>
<p><i>ATM, bad, efficient, Ethernet, external, fiber, include, internal, ISDN, modular, network, private, server, subscriber, swap, system's, user-to-, network-to-, wireless, WLAN, device-, common, own, generic, outgoing, physical, standard</i></p>	<p><i>network interface(s)</i></p>	<p><i>card(s), units, requirements, levels, adapter, enable, PNNI, LNNI, security, supported, connection, SNI, layer, interrupts, CPU, changes, UNI, agreements, LAN, affects,</i></p>
<p><i>access, corresponding, defines, destination, determination, developing, essential, GPRS, higher, implement, IP, ISDN, lower, mapping, multiple, OSI, particular, phase, primary, programmable, range, provisional, standard</i></p>	<p><i>network layer</i></p>	<p><i>protocol(s), multicast, packet, address(es), problem, exceptions, deals, headers, routing, topics, connections, routes, datagrams, solutions, security, comprises, unicast, redundancy, functions, features, multicast, implementations, operations, reachability</i></p>
<p><i>advanced, teleworkers, Alcatel, requirements, security, support, operational, billing, user, cross-connects, assist, automate(d), web-based, Java-based, icon-based, message-based, bilingual, carries, centralized, carrier-class, classical, comprehensive, computers, control, customer, develops, efficient, effective, enterprise, full-featured, IBM, integrated, internet, ISO, leading, layer</i></p>	<p><i>network management</i></p>	<p><i>protocol(s), system(s), software, standpoint, tool(s), applications, -management, products, aspects, details, functions, layers, stations, basics, report, concept, monitors, CNM, CPE, teams, vendors, solutions, procedures, products, platform, architecture, actions, functional, includes, handles, refers, support, capabilities, model, elements, concept, program</i></p>

<i>abstract, backbone, developed, SNA, neuro-fuzzy, including, overall, TCP/IP, oriented, peer-to-peer, OSI, peer, neural</i>	<i>network model</i>	<i>consists, defined, takes, supported, -based,</i>
<i>adjacent, AppleTalk, APPN, associated, mesh, optical, public-, wireless, intermediate, low-entry, ad-hoc, cellular, central</i>	<i>network node(s)</i>	<i>NN, using, determines, starts up, server, interface</i>
<i>correct, day-to-day, encountered, lowering, proper, robust, simplify, wireless, wrap, collocated, synchronous, term, internal, simplifies, provider</i>	<i>network operation(s)</i>	<i>management, situations, centers, expenses, continues, applications, monitoring, group, expertise</i>
<i>address, ATM, core, GSM, home, individual, local, mobile, single, public, visited, access, foreign, human, originating, radio, cellular, CATV, communications, country</i>	<i>network operator(s)</i>	<i>receiving, against, assures, billing, prefers, decides, NO, modulation, investment, launches, starts, hires, needs, service</i>
<i>state-of-art, basic, comprehensive, consolidating, corporate, custom, enterprise, formalized, GPRS, increases, internal, internet, involves, mobile, successful, usage, wireless, interLAN, logical</i>	<i>network security</i>	<i>policy, issues, vendors, system, research, feature, data, device, structure, secrets, solutions, training, threats, measures, management, design, portfolio, arrangements, plays, firm, applies, controls, area, bar, architecture, mechanisms, perimeter</i>
<i>AppleTalk, cost-effective, extend, nonextended, physical, protects, stub, unrepeated, entire, associated, single, local</i>	<i>network segment</i>	<i>distances, length, basis, bandwidth,</i>
<i>in-building, connectionless, dispersed, best-effort, integrated, national, OSI, private, specific, technical, telecommunications, virtual, layer 3, always-on, integrated, managed, public, addresses, area, available, IP-based, broadband, business, Cisco, civic, common, emulating, data, internet global, digital, end-to-end, intelligent, interNAP</i>	<i>network service(s)</i>	<i>allied, CLNS, environments, models, providers, architecture(s), VINES, providers, classes, access, use, protocols, control, delivers, operator, platform, defines, LAP, product, seems, division, corporation, strategy, security</i>
<i>access, agent, broadcast, cable, evaluate, internal, ad hoc, modular, multimode, Xerox, cellular, Orbit, communication, managing, neural, optical,</i>	<i>network systems</i>	<i>delays, partners, design, portfolio, integrity, construction, optimized, manager, management, engineer, Europe, represents, manufacturers, laboratories, suppliers, XNS</i>
<i>distributed, dumbbell, entire, fresh, hypercube, multipoint, needed, optical, physical, predominant, ring, sample, simplistic, specific, tested, star-wired, APPN, current, dynamic, passive-star, ATM, Ethernet, ad hoc, fiber, logical, wireless, large-scale, area</i>	<i>network topology</i>	<i>database(s), map, composed, change(s), environments, information, knowledge, outperformed, grows, remains, independence</i>

2. Familia de *route*. Forma analizada: ROUTER

1. FRECUENCIA.

KEYWORD	Frec. Telec	Frec. Lacell	Ratio	Término, Chung	KEYNESS	P valor
10. ROUTER	3.910	25	0,00011896	ESPECIFICO	11.974,40	0
Miembros relacionados:	F Telec	F Gene	Ratio	Término	Keyness	P valor
ROUTING / Académica	3.542	40	0,00019033	ESPECIFICO	10.690,50	0
ROUTERS	1.891	4	1,90E-05	ESPECIFICO	5.876,00	0
ROUTE / Académica	1.437	1.015	0,00482958	NO	1.655,50	0
ROUTES / Académica	891	419	0,00199369	NO	1.348,40	0
ROUTED / Académica	345	39	0,00018557	NO	848	0
ROUTABLE	32	0	0	inf/esp	100,4	0
REROUTING	22	0	0	inf/esp	69	0
ROUTABILITY	18	0	0	inf/esp	56,5	0
NONROUTABLE	14	0	0	inf/esp	43,9	0
WASTEROUTE	11	0	0	inf/esp	34,5	0
ALLSPFROUTERS	11	0	0	inf/esp	34,5	0
REROUTE	13	3	1,43	NO	26,7	0
ALLDROUTERS	8	0	0	inf/esp	25,1	0,000001
XROUTER	6	0	0	inf/esp	18,8	0,000014
UNROUTABLE	6	0	0	inf/esp	18,8	0,000014
REROUTED	9	3	1,43	NO	16,1	0,000059
ENROUTE	5	0	0	inf/esp	15,7	0,000075
TRACEROUTE	5	0	0	inf/esp	15,7	0,000075
Familia técnica: No, estadística y formalmente.						

2. DISTRIBUCIÓN.

Distribución por áreas	Distribución Keyness	Keykeyword	Nº de textos	Porcentaje
8 (-5)	802	ROUTER	134	8,10%
<i>Dispersion plot</i>				
Area	Words	Hits	Per 1,000	
3 Telematics	1.204.953	2.581	2.14	
082 Esp.Tele	997.674	1.159	1.16	
4 Signal proc.	580.857	120	0.21	
6 Business	373.033	55	0.15	
2 Ar. Comp	329.596	45	0.14	
081 Esp.Sing	867.091	80	0.09	
7 Systems	307.649	19	0.06	
1 Electronics	722.734	33	0.05	

3. COLOCACIÓN.

3.1 Colocados.

N	Collocates	TOTAL	LEFT	RIGHT	L5	L4	L3	L2	L1	*	R1	R2	R3	R4	R5
1	THE	3706	2288	1418	370	278	277	511	852	0	8	464	342	254	350
2	A	1554	981	573	115	108	139	183	436	0	68	179	120	102	104
3	TO	1290	653	637	121	121	192	143	76	0	129	113	138	132	125

4	IS	858	323	535	102	119	66	32	4	0	161	136	67	73	98
5	OF	738	475	263	74	92	119	142	48	0	10	34	30	87	102
6	AND	728	330	398	60	83	80	79	28	0	96	118	47	57	80
7	DESIGNATED	583	463	120	35	28	7	4	389	0	0	12	41	35	32
8	IN	548	278	270	71	57	65	69	16	0	43	55	51	61	60
9	FOR	431	186	245	45	44	34	39	24	0	66	76	26	35	42
10	THAT	423	193	230	23	26	30	76	38	0	66	38	41	32	53
11	NETWORK	343	158	185	32	55	34	25	12	0	8	30	62	41	44
12	AN	330	166	164	22	24	60	60	0	0	1	53	43	29	38
13	THIS	316	150	166	30	27	33	13	47	0	3	58	40	37	28
14	AS	311	141	170	19	32	23	60	7	0	14	45	34	43	34
15	BE	283	141	142	52	48	23	17	1	0	0	42	41	20	39
16	OR	277	129	148	20	23	25	26	35	0	56	24	26	24	18
17	BY	276	199	77	29	33	48	72	17	0	6	12	22	20	17
18	ON	257	135	122	27	22	32	43	11	0	39	14	19	23	27
19	WITH	255	127	128	24	28	34	33	8	0	35	25	19	28	21
20	ID	248	76	172	21	23	22	9	1	0	113	7	18	23	11
21	LINK	243	102	141	34	25	25	15	3	0	15	14	38	44	30
22	LSAS	242	67	175	18	20	19	7	3	0	102	4	25	20	24
23	EACH	238	194	44	8	10	13	23	140	0	0	15	9	13	7
24	IF	235	158	77	26	19	17	70	26	0	6	25	12	15	19
25	S	213	9	204	2	1	2	2	2	0	198	1	1	1	3
26	ARE	210	93	117	53	24	12	4	0	0	13	26	23	30	25
27	BACKUP	208	135	73	17	3	1	114	0	0	0	25	29	13	6
28	FROM	200	131	69	18	20	34	33	26	0	12	9	8	15	25
29	ITS	198	71	127	11	15	4	23	18	0	0	31	44	25	27
30	AREA	192	123	69	23	20	10	36	34	0	0	1	14	22	32
31	LSA	183	53	130	11	18	13	8	3	0	68	2	15	22	23
32	IT	177	49	128	18	19	10	0	2	0	4	44	24	28	28
33	WHEN	177	111	66	16	12	27	47	9	0	10	25	5	7	19
34	HAS	171	41	130	16	11	10	4	0	0	56	34	17	11	12
35	INTERFACE	159	58	101	17	20	17	4	0	0	24	28	18	15	16

3.2 Colocados significativos.

Nº	Colocados	Z	T	MI	Nº	Colocados	Z	T	MI
1	INTERAREA	111,84	3,87	9,71	101	ADVERTISEMENTS	12,15	2,84	4,20
2	ROUTER'S	234,70	15,74	7,80	102	REF	12,74	2,99	4,18
3	REORIGINATE	32,17	2,23	7,71	103	OSPF	47,05	11,09	4,17
4	INELIGIBLE	42,87	3,15	7,54	104	OUTBOUND	10,51	2,50	4,15
5	DESIGNATED	310,89	24,00	7,39	105	REFERENCED	16,84	4,00	4,15
6	SOLICITATION	36,42	3,29	6,94	106	TLV	9,60	2,31	4,11
7	RESTARTING	54,06	5,05	6,84	107	SPF	9,42	2,30	4,06
8	WCCP	41,85	3,96	6,80	108	DIGEST	8,60	2,10	4,06
9	RTX	20,20	2,21	6,39	109	IOS	17,34	4,30	4,02
10	RTR	25,55	2,79	6,39	110	INTERFACE'S	9,24	2,30	4,02
11	RESTARTS	25,55	2,79	6,39	111	ARP	15,20	3,86	3,96
12	ELECTS	20,20	2,21	6,39	112	REACHABILITY	8,23	2,09	3,95
13	NEIGHBOR'S	49,09	5,50	6,32	113	RECEIVES	23,56	5,99	3,95
14	X'S	21,24	2,42	6,27	114	ITSELF	37,70	9,58	3,95
15	BACKUP	122,38	14,23	6,21	115	IGMP	9,00	2,29	3,95
16	NETIRON	22,32	2,61	6,19	116	LSR	16,37	4,18	3,94
17	ORIGINATES	39,90	4,73	6,15	117	ICMP	13,68	3,50	3,94
18	LINKSYS	34,57	4,18	6,10	118	ALIVE	8,14	2,09	3,93
19	ASBR	23,89	2,95	6,03	119	SENDS	26,08	6,79	3,88

20	NSSA'S	17,65	2,20	6,01	120	CISCO'S	11,80	3,09	3,87
21	NSSA	70,84	8,91	5,98	121	PERIODICALLY	13,10	3,48	3,83
22	VRE	19,09	2,41	5,97	122	PE	13,10	3,48	3,83
23	RECOMPUTATION	19,09	2,41	5,97	123	SOUTH	13,99	3,72	3,82
24	RT	52,77	6,67	5,97	124	ADVERTISE	10,39	2,78	3,80
25	LINK'S	33,51	4,29	5,93	125	ABR	12,45	3,35	3,79
26	DISCOVERS	21,69	2,78	5,93	126	SUBNETWORK	11,44	3,08	3,79
27	B'S	16,99	2,20	5,90	127	PROPOSES	9,63	2,62	3,76
28	ELIGIBLE	41,11	5,38	5,87	128	INGRESS	10,75	2,93	3,75
29	BORDER	76,17	10,02	5,85	129	EXTERIOR	8,16	2,26	3,70
30	APPENDS	16,39	2,20	5,80	130	HOP	23,09	6,45	3,68
31	ADVERTISING	60,30	8,15	5,77	131	DROPPING	7,33	2,06	3,66
32	ID	114,11	15,46	5,77	132	TELLS	9,82	2,76	3,66
33	HITLESS	27,10	3,67	5,77	133	HOPS	10,79	3,05	3,65
34	ELECTED	36,16	4,91	5,76	134	INFINIBAND	7,20	2,05	3,62
35	RESTARTED	24,69	3,40	5,72	135	EIGRP	7,20	2,05	3,62
36	RELOADED	15,85	2,19	5,71	136	ARRIVES	10,14	2,90	3,61
37	IDRP	15,85	2,19	5,71	137	DR	17,23	4,94	3,60
38	REDISTRIBUTE	18,50	2,59	5,67	138	PASSIVE	15,99	4,59	3,60
39	ADVERTISES	24,83	3,53	5,63	139	TRANSIT	12,73	3,67	3,59
40	ELIGIBILITY	15,36	2,19	5,62	140	LS	12,70	3,67	3,58
41	FORWARDS	42,02	6,11	5,56	141	FAILS	12,70	3,67	3,58
42	LSAS	104,57	15,23	5,56	142	ACM	7,76	2,24	3,58
43	DECLARING	19,81	2,93	5,51	143	DETERMINES	14,39	4,19	3,56
44	ADVERTISEMENT	36,17	5,36	5,51	144	SHORTEST	14,01	4,09	3,55
45	SOHO	18,46	2,76	5,48	145	BOUNDARY	14,68	4,29	3,55
46	NEIGHBORING	47,30	7,18	5,44	146	RIP	16,71	4,92	3,53
47	STANDBY	25,66	3,91	5,43	147	SECONDARY	17,83	5,25	3,53
48	ORIGINATING	39,95	6,10	5,42	148	DISCOVERY	13,16	3,87	3,53
49	PACKET'S	22,10	3,38	5,42	149	ADJACENT	16,05	4,74	3,52
50	MULTICASTS	16,39	2,58	5,33	150	MULTIPROTOCOL	6,91	2,04	3,52
51	ORIGINATED	40,17	6,39	5,30	151	PROCEEDS	6,85	2,04	3,50
52	ADJACENCIES	31,58	5,06	5,28	152	DEAD	9,16	2,73	3,49
53	LEARNS	18,15	2,92	5,27	153	CONFIGURED	19,78	5,96	3,46
54	DISCARDS	14,82	2,39	5,27	154	CISCO	22,30	6,74	3,45
55	RESTART	33,01	5,33	5,26	155	DESTINED	7,35	2,22	3,45
56	LSA	81,26	13,17	5,25	156	BROADCASTS	8,47	2,57	3,44
57	NBMA	30,52	4,96	5,24	157	ALCATEL	10,36	3,14	3,44
58	ASSIST	38,14	6,23	5,23	158	LSP	9,87	3,01	3,43
59	BR	18,54	3,08	5,18	159	PRECEDENCE	9,39	2,87	3,42
60	ORINATION	14,25	2,38	5,16	160	ASSIGNS	6,64	2,03	3,42
61	HSRP	20,16	3,37	5,16	161	PRIORITY	24,22	7,51	3,38
62	ELECTION	27,68	4,66	5,14	162	NAT	8,19	2,55	3,36
63	ORIGINATE	25,02	4,23	5,13	163	ENDPOINT	7,57	2,38	3,33
64	UBR	15,07	2,57	5,11	164	ROUTERS	30,53	9,62	3,33
65	ADVERTISED	28,96	4,95	5,10	165	AUTONOMOUS	11,04	3,49	3,33
66	TRANSLATOR	17,93	3,07	5,09	166	BGP	13,55	4,31	3,31
67	DSU	13,74	2,38	5,06	167	DEFAULT	19,86	6,35	3,29
68	LER	17,92	3,21	4,96	168	EXCEEDED	6,86	2,20	3,28
69	INITIALIZES	12,04	2,16	4,95	169	DETECTS	6,25	2,01	3,28
70	PRI	15,08	2,74	4,93	170	PASSES	9,61	3,10	3,26
71	KNOWS	30,91	5,64	4,91	171	HORIZON	7,34	2,37	3,26
72	REACHABLE	19,16	3,61	4,82	172	FORWARDED	9,59	3,10	3,26
73	PICTURED	11,38	2,16	4,80	173	INTER	15,89	5,14	3,25

74	NVRAM	11,38	2,16	4,80	174	DHCP	8,69	2,83	3,24
75	HELLO	44,93	8,57	4,78	175	OPAQUE	8,63	2,82	3,22
76	STUB	28,39	5,45	4,76	176	INTERFACES	22,13	7,25	3,22
77	JUNIPER	11,18	2,15	4,75	177	VRF	6,66	2,19	3,22
78	NETWORK'S	17,98	3,47	4,75	178	INFORM	6,66	2,19	3,22
79	PREFIXES	20,87	4,08	4,71	179	RFC	11,71	3,88	3,19
80	IDS	19,30	3,84	4,66	180	REPLY	8,88	2,95	3,18
81	HELPER	11,78	2,35	4,65	181	FORWARDING	16,04	5,34	3,18
82	CALCULATING	20,40	4,07	4,65	182	CE	8,45	2,81	3,18
83	REPLIES	14,30	2,88	4,62	183	HUB	9,53	3,20	3,15
84	SLIP	10,64	2,15	4,62	184	LABEL	17,07	5,75	3,14
85	TREATS	12,39	2,53	4,58	185	FLOODING	9,11	3,07	3,14
86	ATTACHED	42,01	8,67	4,55	186	DESTINATION	21,84	7,36	3,14
87	INTRA	26,17	5,41	4,55	187	NORTH	10,76	3,65	3,12
88	NLSP	11,32	2,34	4,54	188	DECLARED	9,38	3,19	3,11
89	BOOT	18,88	3,94	4,52	189	BIDIRECTIONAL	6,37	2,17	3,11
90	MLINK	11,99	2,53	4,49	190	UPGRADES	6,34	2,16	3,10
91	DIJKSTRA	10,01	2,13	4,46	191	INTERNETWORK	8,18	2,79	3,10
92	LSA'S	14,06	3,02	4,44	192	SUBNET	12,08	4,14	3,09
93	CALCULATES	14,26	3,15	4,35	193	ROUTE	21,96	7,54	3,09
94	ADJACENCY	17,19	3,80	4,35	194	FAILED	8,13	2,79	3,09
95	LISTING	15,90	3,55	4,32	195	HIGHEST	13,30	4,58	3,08
96	ASSISTING	10,20	2,32	4,27	196	CONNECTS	9,20	3,18	3,07
97	NEIGHBORS	28,52	6,50	4,27	197	PACKETS	26,97	9,38	3,05
98	PREFIX	24,13	5,53	4,25	198	NEWLY	6,61	2,32	3,02
99	GRACE	10,09	2,32	4,24	199	LIKEWISE	6,56	2,32	3,00
100	NEIGHBOR	33,72	7,75	4,24					

3.3 Definición del nodo en el diccionario.

1. El *Advanced Learner's Dictionary* de Cambridge (2003) no registra la definición de *router*. El lema más cercano a *router* en este diccionario es *route*:

Route (noun):

1. *A particular way or direction between places.*
2. *A method of achieving something.*
3. *A set of regular visits that you make to a number of places or people, especially in order to deliver products as part of your job.*

2. Definición de *router* en el diccionario especializado Webster:

Router (noun):

1. *A worker who routes shipments for distribution and delivery.*
2. *(computer science) a device that forwards data packets between computer networks.*
3. *A power tool with a shaped cutter; used in carpentry for cutting grooves.*

Specialty Definition:

Router (noun): *A device which forwards packets between networks. The forwarding decision is based on network layer information and routing tables, often constructed by routing protocols.*

A router consists of a computer networking device that determines the next network point to which to forward a data packet toward its destination, a process known as routing. Routing occurs at layer 3 of the OSI seven-layer model.

3.4 Colocados contiguos.

a) Colocado + router.		
Total colocados seleccionados: 39	Frecuencia máxima: 389	Frecuencia mínima: 7
<ul style="list-style-type: none"> - Sustantivos, total 15: <i>border 79, edge 22, default 21, north 17, boundary 17, south 16, Cisco 16, access 14, hop 12, core 11, root 10, switch 9, broadband 9, gateway 8, center 8.</i> - Adjetivos, total 12: <i>conventional 19, highest 18, virtual 18, local 16, traditional 15, new 11, passive 11, active 10, wireless 9, first 9, multiple 8, capable 8.</i> - Participios verbales que actúan como adjetivos, total 9: <i>designated 389, advertising 40, neighboring 33, restarting 22, originating 17, calculating 11, enabled 10, switching 9, filtering 7.</i> - Siglas, total 3: <i>OSPF 43, IP 28, ICMP 8.</i> 		
b) Router + colocado.		
Total colocados seleccionados: 27	Frecuencia máxima: 42	Frecuencia mínima: 7
<ul style="list-style-type: none"> - Sustantivos, total 11: <i>priority 42, vendors 23, configuration 15, discovery 14, architectures 14, election 13, interfaces 12, functions 11, field 11, advertisement 9, software 7, implementations 7.</i> - Participios verbales que actúan como adjetivos, total 1: <i>based 10.</i> - Formas personales de verbos, total 10: <i>assist 27, receives 21, sends 20, uses 11, originates 9, maintains 8, determines 8, calculates 7, performs 7, constructs 7.</i> - Siglas, total 5: <i>LSA S 102/LSA 68, D 14, IDS 10, LSR 9.</i> 		

3.5 Identificación de colocados contiguos especializados en las líneas de concordancia.

GPRS and 802.11b networks. Instead of only relying on Router Advertisement to detect a link change, attacks resulting from repetitive origination of the same router advertisement or origination of a large number of advertisements can cause a link change. The MN can use the Link-Type indication associated to each Router Advertisement (thus to each Care-of-Address Request and Reply, Redirect, Time Exceeded, and Router Advertisement and Router Solicitation. If an ICMP Router-Discovery Protocol (IDRP). IDRP uses Router-Advertisement and Router-Solicitation messages it require manual configuration by an administrator. Router-Advertisement messages enable hosts to discover configured in a way to avoid sending unsolicited router advertisement messages or sending them at a frequency attached subnets. Each router periodically multicasts Router-Advertisement messages from each of its interfaces. IPv6-enabled routers can advertise prefixes in router advertisement messages, as well as function as default routers. Note that the MN does not need to wait for the periodic Router Advertisement, but explicitly solicits it without having to wait for the next scheduled router advertisement. IPv6 prefixes for identifying default routers for the link. Router advertisements are transmitted occasionally and 11.5 of [3]), that use the advertisement interval option of Router Advertisements (RA) and Neighbor Solicitation (NS) messages. The 64-bit prefixes (see next paragraph) included in the router advertisements. These subsequent 128-bit

Concordancias de router advertisement.

a) Colocado + router:

<i>access router</i>	default router	originating router
<i>active router</i>	designated router	OSPF router
advertising router	<i>edge router</i>	passive router
border router	<i>enabled router</i>	restarting router
boundary router	<i>filtering router</i>	<i>root router</i>
<i>broadband router</i>	<i>gateway router</i>	<i>single router</i>
calculating router	hop router	<i>switch router</i>
<i>capable router</i>	ICMP router	<i>traditional router</i>
<i>center router</i>	<i>IP router</i>	<i>virtual router</i>
Cisco router	<i>local router</i>	<i>wireless router</i>
<i>conventional router</i>	<i>multiple router</i>	
<i>core router</i>	neighboring router	

b) Router + colocado:		
router advertisement	router <i>field</i>	router LSA(s)
router <i>architectures</i>	router <i>functions</i>	router LSR
router assist	router IDS	router priority
router <i>configuration</i>	router <i>implementations</i>	router <i>software</i>
router election	router interfaces	router <i>vendors</i>

4. AGRUPACIONES LÉXICAS O CLUSTERS.

4.1 Distribución de clusters.

Combinaciones	Volumen	Frecuencia máxima
2-cluster	449	92
3-cluster	150	167
4-cluster	161	52
5-cluster	180	12
6-cluster	207	8

4.2 Inclusión de combinaciones especializadas en unidades más amplias.

<p>advertising router referenced advertising router the advertising router field</p> <p>border router NSSA border router area border router</p> <p>root router domain root router</p> <p>center router data center router</p> <p>router field the advertising router field</p> <p>router assist router assist mechanism simple router assist mechanism router assisting control tree router assisting control tree configuration</p>	<p>designated router backup designated router designated router election elected designated router</p> <p>switch router label switch router label switch router LSR</p> <p>OSPF router OSPF router ID a router's OSPF router</p> <p>router election designated router election backup designated router election designated router election algorithm</p> <p>router LSR label switch router LSR</p> <p>router vendor traditional router vendors</p>
--	---

4.3 Colocados contiguos de las combinaciones especializadas.

pre +	COMBINACIÓN	+ post
LSA's	advertising router	<i>field(s)</i>
<i>type 5, area, AS, attains, NNSA, real</i>	border router	translator, advertises, ABR, uses, summarizes, routing, originates, perform
<i>Data</i>	<i>center router</i>	
<i>highlighted, a number of, problems, today's</i>	<i>conventional router</i>	VRF, vendors, architectures, management
<i>level 1, backup, be, become router's, configured, current, DR, elected, link's, neighbor's, L's, network's, new, potential, present, previous</i>	designated router	election, elected, priority, ID, field, performs, changes, becomes, builds, originates, is flooding, eligibility, smoother
destination, particular, router's, specific	OSPF router	<i>dynamically, computes, converses, finds, ID, collects, calculates, maintains, interface, vendors</i>
<i>domain</i>	<i>root router</i>	
label	<i>switch router</i>	LSR, edge, segments
HSRP, single, standardized, track, VR, VRE	<i>virtual router</i>	<i>IP, definitions, redundancy, engine, VR, VRE</i>
intra-, wavelength, traditional, conventional,	<i>router architecture(s)</i>	<i>challenges, employ, face, rely, SPs</i>
<i>simple, design, proposed, PGM, generic</i>	router assist	<i>mechanism, supports, GRA</i>

<i>switch, network, Cisco, additional, simplest, following, downstream</i>	<i>router configuration</i>	<i>command, excerpt, mode, possible, required, specifies, using, refer</i>
designated	router election	<i>algorithm, process</i>
advertising, designated	<i>router field</i>	<i>appears</i>
<i>neighbour, highest, designated, router's, IPv6, possible, own, new, OSPF, destination, corresponding, higher</i>	router ID	<i>allows, appearing, assignment, field, found, takes, is chosen, is seen, is changed, specified, takes, identifies</i>
<i>Cisco, designated, configurable, outgoing, separate, OSPF, single, functioning, certain, looped, generally</i>	router interface(s)	<i>basis, connects, depend, fails, failures, information, IDs, parameters, composing, excepting</i>
<i>RTX's, separate, inter-area-, pre-start, NSSA's, router's, above, single</i>	router LSR	<i>carries, contains, describes, indicates, originated, option(s)</i>
<i>traditional, conventional, PE</i>	<i>router vendors</i>	<i>often, provide</i>

3. Familia de layer. Forma analizada: LAYER

1. FRECUENCIA.

KEYWORD	Frec. Telec	Frec. Lacell	Ratio	Término, Chung	KEYNESS	P valor
15. LAYER / Académica	4.425	569	29,5353086	NO	10.604,80	0
Miembros relacionados:	F Telec	F Gene	Ratio	Término	Keyness	P valor
LAYERS / Académica	839	374	8,51982869	NO	1.307,70	0
SUBLAYER	112	0	inf/esp	inf/esp	351,3	0
LAYERED / Académica	163	86	7,19829527	NO	230,5	0
MULTILAYER	49	1	186,095805	ESPECIFICO	144,3	0
SUBLAYERS	28	0	inf/esp	inf/esp	87,8	0
PRIMELAYER	15	0	inf/esp	inf/esp	47	0
MONOLAYERS	17	1	64,5638508	ESPECIFICO	46,1	0
MONOLAYER	25	8	11,8683549	NO	45,6	0
MULTILAYERED	20	5	15,1914943	NO	40	0
MULTILAYERS	5	0	inf/esp	inf/esp	15,7	0,000075
UPPERLAYER	5	0	inf/esp	inf/esp	15,7	0,000075
INTERLAYER	5	0	inf/esp	inf/esp	15,7	0,000075

Familia técnica: No, estadística y formalmente.

2. DISTRIBUCIÓN.

Distribución por áreas	Distribución Keyness	Keykeyword	Nº de textos	Porcentaje
9	082	LAYER	241	14,57 %
<i>Dispersion plot</i>				
Area	Words	Hits	Per 1,000	
082 Esp.Tele	997.694	1.960	1.96	
1 Electronics	722.798	587	0.81	
3 Telematics	1.204.957	816	0.68	
081 Esp.Sing	867.140	566	0.65	
5 Materials	101.237	52	0.51	
7 Systems	307.660	122	0.40	
4 Signal proc.	580.874	229	0.39	
2 Ar. Comp	329.602	46	0.14	
6 Business	373.038	26	0.07	

3. COLOCACIÓN.

3.1 Colocados.

N	Collocates	TOTAL	LEFT	RIGHT	L5	L4	L3	L2	L1	*	R1	R2	R3	R4	R5
1	THE	4056	2481	1575	332	269	525	1193	162	0	7	524	380	333	331
2	OF	1158	626	532	114	150	207	112	43	0	150	60	77	120	125
3	AND	1099	481	618	109	117	97	117	41	0	108	159	103	115	133
4	A	1033	627	406	72	115	148	194	98	0	1	122	92	95	96
5	TO	926	489	437	88	110	175	91	25	0	54	83	101	94	105
6	IS	847	296	551	125	69	71	30	1	0	167	99	111	84	90
7	IN	665	331	334	74	54	141	49	13	0	66	74	81	58	55
8	NETWORK	623	478	145	33	38	24	83	300	0	2	24	44	40	35
9	FOR	450	215	235	26	47	68	53	21	0	27	56	51	46	55

10	PHYSICAL	432	374	58	20	11	11	17	315	0	0	8	16	19	15
11	DATA	415	277	138	35	26	32	183	1	0	21	15	27	34	41
12	LINK	350	293	57	19	8	10	12	244	0	0	5	9	28	15
13	PROTOCOL	348	99	249	42	28	11	8	10	0	126	30	31	24	38
14	THAT	313	132	181	29	43	27	19	14	0	22	51	38	36	34
15	AS	307	121	186	37	32	30	16	6	0	13	38	51	47	37
16	WITH	303	171	132	29	53	44	25	20	0	17	30	31	32	22
17	AT	301	234	67	21	42	110	19	42	0	10	11	15	12	19
18	TRANSPORT	295	246	49	8	15	5	8	210	0	4	8	17	12	8
19	THIS	288	149	139	22	22	16	11	78	0	0	52	27	26	34
20	ARE	252	92	160	32	23	25	10	2	0	20	37	34	38	31
21	ON	249	141	108	32	18	40	32	19	0	20	20	18	23	27
22	APPLICATION	243	194	49	14	15	4	11	150	0	0	6	20	15	8
23	OSI	235	137	98	20	11	24	41	41	0	7	8	35	30	18
24	AN	230	132	98	16	21	23	72	0	0	1	19	18	24	36
25	BY	230	129	101	21	25	53	23	7	0	6	17	25	27	26
26	PROTOCOLS	230	45	185	17	8	10	10	0	0	125	23	16	3	18
27	UPPER	223	199	24	5	5	9	4	176	0	0	4	4	8	8
28	BE	187	69	118	37	23	7	2	0	0	0	28	31	25	34
29	OR	178	90	88	16	17	20	23	14	0	10	26	20	15	17
30	FROM	167	94	73	9	16	47	14	8	0	29	10	13	9	12
31	SERVICES	165	59	106	9	23	14	8	5	0	18	30	23	13	22
32	CONTROL	160	63	97	11	14	11	11	16	0	4	40	13	20	20
33	ACCESS	155	105	50	15	14	16	7	53	0	1	5	12	19	13
34	CAN	155	37	118	21	6	8	2	0	0	35	22	21	20	20
35	IT	150	48	102	21	18	4	5	0	0	0	25	36	20	21

3.2 Colocados significativos.

Nº	Colocados	Z	T	MI	Nº	Colocados	Z	T	MI
1	WEIGHTS	9,63	3,10	3,27	54	DIOXIDE	16,89	3,06	4,93
2	MAC	34,77	9,21	3,83	55	PLANES	10,51	3,04	3,58
3	SWITCHES	30,97	9,17	3,51	56	PHYSICAL	105,04	20,03	4,78
4	SESSION	30,53	8,55	3,67	57	BACKPLANE	9,59	2,99	3,36
5	LAYERS	28,66	8,25	3,59	58	ABSORBED	11,87	2,97	4,00
6	HEADER	27,46	8,20	3,49	59	ROUGHLY	9,06	2,96	3,23
7	PRESENTATION	39,74	8,18	4,56	60	SCL	11,46	2,95	3,91
8	OXIDE	37,87	7,56	4,65	61	COMMUNICATES	11,46	2,95	3,91
9	THICKNESS	36,67	7,25	4,68	62	RR	11,08	2,94	3,83
10	ADAPTATION	30,35	6,72	4,35	63	TUNNELING	8,59	2,93	3,10
11	VPNS	23,51	6,47	3,72	64	IPX	8,43	2,92	3,06
12	APPLETALK	33,18	6,45	4,72	65	SERVERIRON	15,56	2,90	4,85
13	PLANE	18,86	6,23	3,19	66	POLYSILICON	15,35	2,89	4,81
14	SI	17,18	5,70	3,18	67	DNA	14,95	2,89	4,74
15	HEADERS	26,86	5,65	4,50	68	SIO	36,30	2,81	7,38
16	SUBLAYER	29,70	5,22	5,02	69	GROWN	8,25	2,80	3,12
17	RESPONSIBLE	14,94	5,20	3,05	70	ANONYMITY	8,25	2,80	3,12
18	LSW	54,64	5,15	6,81	71	PHOTONS	10,61	2,79	3,85
19	THIN	14,49	4,66	3,27	72	SUBLAYERS	16,43	2,75	5,16
20	POLY	20,59	4,65	4,29	73	QPSK	7,88	2,66	3,14
21	SUBSTRATE	13,59	4,60	3,13	74	EOMPLS	9,63	2,62	3,76
22	PASSED	13,02	4,56	3,03	75	CLADDING	9,14	2,60	3,63
23	DEFECT	15,68	4,42	3,66	76	DOCSIS	8,70	2,58	3,51
24	PHY	29,73	4,27	5,60	77	CMT	8,00	2,54	3,31
25	TRAILER	22,31	4,10	4,89	78	DDP	10,94	2,51	4,25

26	DECNET	16,77	4,10	4,06	79	TRANSPARENCY	7,30	2,50	3,10
27	MANAGES	15,67	4,07	3,89	80	MULTILAYER	10,56	2,50	4,16
28	ENCAPSULATED	15,95	4,07	3,94	81	TCO	10,10	2,49	4,05
29	PPP	12,70	4,02	3,32	82	TIN	9,60	2,47	3,92
30	DEPLETION	15,95	3,98	4,01	83	EIA	9,24	2,46	3,82
31	DEPOSITED	15,22	3,96	3,89	84	NEURONS	9,07	2,45	3,77
32	INNER	14,22	3,92	3,72	85	PHOTOCONDUCTIVE	19,06	2,41	5,97
33	THICKNESSES	25,77	3,91	5,44	86	NANOPARTICLE	19,06	2,41	5,97
34	OUTER	12,48	3,75	3,47	87	PLP	16,42	2,40	5,55
35	LAYERED	12,87	3,67	3,62	8	ACKNOWLEDGMENTS	7,56	2,38	3,33
36	NETWARE	14,76	3,64	4,04	89	RESTORATION	7,40	2,37	3,28
37	DATAGRAM	12,10	3,64	3,47	90	VLANS	6,93	2,34	3,13
38	DEMODULATOR	20,65	3,63	5,02	91	LI	6,85	2,34	3,10
39	MULTIHOMING	19,88	3,62	4,92	92	BURIED	10,49	2,33	4,34
40	ISO	11,75	3,62	3,40	93	OAM	9,29	2,30	4,03
41	EPITAXIAL	20,06	3,62	4,94	94	ENCAPSULATING	8,96	2,29	3,94
42	HANDLES	11,39	3,60	3,32	95	ETCHED	7,60	2,24	3,53
43	PDU	11,91	3,53	3,51	96	TERMINATES	7,40	2,23	3,46
44	THICKER	23,03	3,52	5,42	97	CDP	7,27	2,22	3,42
45	SSL	11,60	3,42	3,53	98	NITRIDE	6,70	2,19	3,23
46	ABSORBANCE	28,87	3,42	6,16	99	CU	6,36	2,17	3,11
47	CAP	14,66	3,41	4,21	101	LINK	57,52	17,06	3,51
48	THICK	12,79	3,36	3,86	102	TRANSPORT	92,31	16,62	4,95
49	VINES	17,60	3,34	4,80	103	OSI	101,96	15,00	5,53
50	DRAG	16,06	3,32	4,55	104	UPPER	99,80	14,61	5,54
51	SOCKETS	14,51	3,29	4,28	105	PROTOCOLS	53,55	14,11	3,85
52	SOCKET	9,86	3,12	3,32	106	ATM	34,50	10,63	3,40
53	THINNER	18,37	3,07	5,16	107	LOWER	31,89	10,48	3,21

3.3 Definición del nodo en el diccionario.

- Definición de *layer* en el *Advanced Learner's Dictionary* de Cambridge (2003):

Layer (noun): *a level of material, such as a type of rock or gas, which is different from the material above or below it, or a thin sheet of a substance: The ozone layer. A thick layer of clay lies over the sandstone.*

- Definición de *layer* en el diccionario especializado Webster:

Layer (noun):

- Single thickness of usually some homogeneous substance.*
- A relatively thin sheet like expanse or region lying over or under another.*
- A hen that lays eggs.*
- Thin structure composed of a single thickness of cells.*

Specialty Definition:

In telecommunications, the term layer has the following meanings:

- In radio wave propagation, see F region.*
- In telecommunications networks and open systems architecture, a group of related functions that are performed in a given level in a hierarchy of groups of related functions.*

Note: In specifying the functions for a given layer, the assumption is made that the specified functions for the layers below are performed, except for the lowest layer.

3.4 Colocados contiguos.

a) Colocado + layer.		
Total colocados seleccionados: 58	Frecuencia máxima: 315	Frecuencia mínima: 6
<ul style="list-style-type: none"> - Sustantivos, total 25: <i>network 300, link 244, transport 210, application 150, access 53, session 49, oxide 46, presentation 44, adaptation 27, cross 26, management 24, depletion 18, internet 18, charge 12, boundary 11, defect 17, buffer 15, sockets 11/socket 8, user 11, distribution 10, output 9, polysilicon 8, backplane 7, nanoparticle 6.</i> - Adjetivos, total 24: <i>physical 315, upper 176, lower 102/lowest 7, optical 66, higher 60, first 20, second 19, single 19, seven 15, one 13, poly 13, two 11, outer 10, epitaxial 12, traditional 10, new 10, pure 7, next 6, photoconductive 6, three 6, dielectric 11, metal 11, logic 8.</i> - Participios verbales que actúan como adjetivos, total 1: <i>cladding 6.</i> - Participios verbales, total 1: <i>including 8.</i> - Siglas, total 7: <i>MAC 52, ATM 50, IP 21, MPLS 15, CAP 12, RR 10, SONET 8.</i> 		
b) Layer + colocado.		
Total colocados seleccionados: 37	Frecuencia máxima: 126	Frecuencia mínima: 6
<ul style="list-style-type: none"> - Sustantivos, total 26: <i>protocol 126/protocols 125, signal 65/signals 7, addresses 25/address 23, functions 24, security 23, parameters 12, implementations 11/implementation 10, entities 11, tree 10, uses 9, processes 9, device 9, technology 8, entity 8, design 8, configuration 8, frames 7, controller 7, acknowledgments 7, module 6, deposition 6, characteristics 6.</i> - Adjetivos, total 2: <i>multicast 13, normal 10.</i> - Formas personales de verbos, total 5: <i>creates 11, manages 9, defines 9, corresponds 8, represents 7.</i> - Siglas, total 4: <i>SSL 11, AAL 7, LI 7, SCL 6.</i> 		

3.5 Identificación de colocados contiguos especializados en las líneas de concordancia.

kdrop to examine the functions of the data link layer (Layer 2). Data Link Function 1: Arbitrarily set its MSS value to the MTU at the Data Link layer minus 40 bytes for the IP and TCP header pattern into a buffer. It notifies the data link layer that a frame has been received after decoding whose source and destination are data link layer entities. A frame is composed of the data link layer. Across the Um interface, the data link layer is a modified version of the LAPD protocol. Upper-Layer Entities Makes Up the Data Link Layer Frame. A packet is an information unit, local termination reduces the likelihood of link layer timeouts across WANs. Similarly, DLSw over the air interface. Layer 2 is the data link layer. Across the Um interface, the data link layer connections helps to reduce the likelihood of link layer timeouts across WANs. The local mapping Node Addresses. LLAP acquires data link layer node addresses dynamically. The process of mapping this address onto data-link layer multicast addresses where they exist. Local media and dynamically acquiring data link layer node addresses. Regulating Node Access. TokenTalk. TokenTalk extends the data link layer to allow the AppleTalk protocol suite to operate over a network interface at OSI layer 2, (the data link layer). An example would be a PBX that terminates traffic at their destination network. Data Link Layer Functions and Protocols. The data link layer entities is encapsulated in the data link layer header and trailer. Figure 1-9 illustrates how to connect more seamlessly across heterogeneous link layer technologies, and working closely with higher layers. SNA and NetBIOS traffic rely on link layer acknowledgments and keepalive message. Local termination eliminates the requirement for link layer acknowledgments and keepalive message formation intended for the destination data link layer entity and is typically 5 bytes long. The reference model: the physical layer, the data link layer, and the network layer. Figure 3-1 illustrates the reference model. A frame is composed of the data link layer header (and possibly a trailer) and upper-

Concordancias de link layer.

a) Colocado + layer:

<i>access layer</i>	epitaxial layer	<i>output layer</i>
adaptation layer	<i>internet layer</i>	oxide layer
<i>application layer</i>	<i>IP layer</i>	phy layer
ATM layer	link layer	
backplane layer	<i>logic layer</i>	physical layer
<i>boundary layer</i>	MAC layer	photoconductive layer
<i>buffer layer</i>	<i>MPLS layer</i>	<i>poly layer</i>
<i>charge layer</i>	<i>management layer</i>	polysilicon layer
depletion layer	<i>network layer</i>	presentation layer
defect layer	<i>optical layer</i>	<i>session layer</i>
<i>dielectric layer</i>	OSI layer	RR layer

b) Layer + colocado:

<i>layer configuration</i>	<i>layer frame</i>	<i>layer protocol</i>
<i>layer controller</i>	<i>layer functions</i>	<i>layer SCL</i>
<i>layer deposition</i>	<i>layer implementation</i>	<i>layer security</i>
<i>layer design</i>	<i>layer multicast</i>	<i>layer signal</i>
<i>layer device</i>	<i>layer parameters</i>	<i>layer SSL</i>
<i>layer entity</i>	<i>layer processes</i>	<i>layer tree</i>

4. AGRUPACIONES LÉXICAS O CLUSTERS.

4.1 Distribución de clusters.

Combinaciones	Volumen	Frecuencia máxima
2-cluster	330	1.124
3-cluster	738	186
4-cluster	517	101
5-cluster	246	43
6-cluster	143	12

4.2 Inclusión de combinaciones especializadas en unidades más amplias.

<p>access layer <i>network access layer</i> <i>data access layer</i> adaptation layer <i>ATM adaptation layer</i> <i>adaptation layer AAL</i> <i>ATM adaptation layer AAL</i> application layer <i>network application layer</i> <i>application layer attacks</i> <i>OSI application layer</i> <i>application layer data</i> <i>application layer features</i> <i>application layer services</i> <i>application layer functions</i> <i>application layer protocol(s)</i> <i>netware application layer</i> <i>netware application layer services</i> logic layer <i>business logic layer</i> MAC layer <i>MAC layer module</i> <i>MAC-layer address</i> management layer <i>network management layer</i> <i>element management layer</i> <i>request management layer</i> <i>service management layer</i> <i>product request management layer</i> network layer <i>network layer protocol(s)</i></p>	<p>boundary layer <i>planetary boundary layer</i> <i>boundary layer normal pressure</i> <i>boundary layer normal pressure drag</i> buffer layer <i>dioxide buffer layer</i> <i>silicon dioxide buffer layer</i> depletion layer <i>depletion layer capacitance</i> <i>charge depletion layer</i> <i>outside the depletion layer</i> <i>space charge depletion layer</i> link layer <i>data link layer</i> <i>link layer protocol</i> <i>link layer acknowledgements</i> <i>link layer multicast</i> <i>PPP link layer</i> <i>link layer address(es)</i> <i>link layer frame(s)</i> <i>link layer manages</i> <i>link layer header</i> <i>link layer implementation</i> <i>link layer security</i> <i>link layer technology</i> <i>OSI data link layer</i> <i>data link layer addresses</i> <i>PPP link layer frames</i> <i>data link layer multicast</i> <i>data link layer manages</i> <i>link layer multicast addresses</i></p>
---	---

<p>network layer address(es) OSI network layer destination network layer multiple network layer network layer multicast network layer routing network layer addressing network layer functions network layer header network layer forwarding IP network layer network layer tree primary network layer model network layer network layer entities network layer packet protocol network layer multiple network layer protocols destination network layer address primary network layer routing network layer routing protocol network layer protocol configuration network layer multicast addresses network layer protocol type presentation layer model presentation layer OSI presentation layer presentation layer implementation session layer session layer protocol OSI session layer session layer implementation optical layer WDM optical layer reconfigurable optical layer layer controller backplane- layer controller layer controller family physical layer controller backplane- layer controller family backplane physical- layer controller layer design cross layer design layer function transport layer functions network layer functions application layer functions OSI transport layer functions layer entity transport layer entity network layer entities upper-layer entities layer frame link layer frame(s) physical layer frame PPP link layer frames ISDN link layer frame layer SCL charge layer SCL space charge layer SCL layer signal upper- layer signal lower- layer signal layer signal power upper- layer signal power lower- layer signal power</p>	<p>link layer multicast data link layer protocol(s) data link layer implementation data link layer frame data link layer addresses physical layer physical layer parameters ISDN physical layer physical layer configuration ATM physical layer physical layer device physical layer specifications physical layer wireless physical layer implementations physical layer manages physical layer controllers docsis physical layer OSI physical layer physical layer adaptation physical layer characteristics physical layer configurations physical layer extension physical layer frame physical layer functions physical layer information physical layer interface physical layer optimization physical layer phy physical layer wireless security ISDN physical layer frame layer SSL socket layer SSL secure layer SSL secure socket layer SSL layer protocol network layer protocol(s) upper-layer protocol(s) transport layer protocol(s) higher-layer protocols link layer protocol lower-layer protocol(s) session layer protocol cross-layer protocol layer protocol configuration packet-layer protocol application layer protocol layer protocol engineering layer protocol information layer protocol type multiple network layer protocol track transport layer protocol network layer protocol configuration network layer protocol negotiation cross-layer protocol engineering standard track transport layer protocol network layer protocol configuration negotiation layer security transport layer security layer security features layer security TLS link layer security transport layer security TLS layer tree transport layer tree network layer tree</p>
--	--

4.3 Colocados contiguos de las combinaciones especializadas.

pre +	COMBINACIÓN	+ post
<i>data, physical, network, medium, wireless</i>	<i>access layer</i>	<i>provides, via, software, manages, roughly, connections, discuss, really, includes, defines, relates</i>
<i>ATM, efficient</i>	<i>adaptation layer</i>	<i>AAL1, resides, ATM, SEAL, AAL, SAAL, software, prepares, converts, used</i>
<i>OSI, network, increases, TCP/IP, full, applied to, carry, application-layer-to-, improve, standard, sensitive, web, uses, netware, model, protocols, proprietary, receiving, performs, well-known</i>	<i>application layer</i>	<i>adaptation, assists, attacks, building, calls, code, complexity, correlates, corresponds, creates, data, data-access, determines, application, features, functions, implementation(s), include(s), interacts, issues, processes, protocol(s), layer 7, message(s), provides, processes, recovery, security, services, statistics, support, gateways, path</i>
<i>combines, unassigned, good, PI-, rate-based</i>	<i>ATM layer</i>	<i>7, 2, cell(s), -combined, contain, performance, places, processor, protocol, provides, Qos, question, rate, transfers</i>
<i>measuring, entirely, show, planetary</i>	<i>boundary layer</i>	<i>normal, effects</i>
<i>nitride, index, dioxide, isolating, substrate</i>	<i>buffer layer</i>	<i>below, 250 nm</i>
<i>charge</i>	<i>depletion layer</i>	<i>barrier, capacitance</i>
<i>data, rely on, provides, sender's, PPP, SNAP, radio, OSI, minimum, dialup, SNA, first, popular, various, different, hardware, heterogeneous, traditional, WDM, WLAN</i>	<i>link layer</i>	<i>3, acknowledgements, acks, uses, paging, address(es), casts, code, communication, component, comprised, contains, creates, solutions, defined, device(s), encapsulation, entity(ies), events, error, feature(s), frame(s), framing, functions, header, helps, identifies, implementation(s), includes, information, 2, load, MAC, manages, medium, multicast, node, network, proprietary, information, security, signalling, specifications, technology(ies), timeouts, transparent, integrity</i>
<i>business, presentation</i>	<i>logic layer</i>	<i>reside</i>
<i>scheduling, real, pre-packaged, new, control</i>	<i>MAC layer</i>	<i>appears, controls, entity, features, function, messages, module, protocol, scheduling, security, address(es), bridge(s), performance, switching, technology, -to-IP</i>
<i>request, service, network, communication, element, business, mobility</i>	<i>management layer</i>	<i>CM, element, implementing, MM, RADview, represents, service</i>
<i>destination, protocol, provisional, VINES, OSI, NSAP, mapping, three, defines, computation, IPv6, special, GPRS, IP, access, programmable, host-to-, ISDN, A-PLP, netware, destination, protocol, single, optimise, using, particular, essential, XNS, multiple, chosen, different, standard, bring up, underlying, primary, employ, provides, model, lower, higher, perform(s), developing, implantation</i>	<i>network layer</i>	<i>address(es), groupings, adds, allows, comprises, connections, internet, creates, datagram, deals, defines, -dependent, takes, drivers, type, solutions, encapsulates, entity(ies), exceptions, features, forwarding, functions, header(s), implementations, -independent, IONL, IP, LAPB, layer 3, manages, multicast, needs, unicast, operation(s), packet, tasks, protocol(s), provides, reachability, redundancy, represents, routing,</i>

		security, service(s), tree, specification
WDM, new, static, topology, reconfigurable,	optical layer	defined, do, has, packet, protection, provides, control, survivability, virtual, restoration , technology
underlying, basic, perfect, COFDM, fibrechannel, cell, DOCSIS , optimum, efficient, frame, discrete, upstream, return-path, ISO , OSI , view, ISDN, vehicles, sequence of, essential, zero, ATM , ideal, different, particular, WAN, Ethernet, exploiting, popular, adding, high-speed, NGSS, ACM, TSMA, excluding, MAC , full, link , absolute, main, BRI, PRI, three, infrared, separate, future, efficient, model, receiving, adaptive, evaluate, security-, several, proprietary, refused, multipair, cutting-edge, backplane	physical layer	allows, simulation, according, achieving, activities, ANSI, adaptability, adaptation , based, adaptive, -analogous, chips, transmission, assumptions, attributes, characteristics, choice, compatibility, component, configuration(s), connection, contains, convergency, convolutional, defines, delay, dependent, device, differences, encodes, parameters, error, extension, especially, failure, frame(s), framework, functions, LAN, granularity, hardware, implementation(s), information, interface(s), manages , operating, operations, optimization, PHY , overhead, passes, permits, PL, routing, problems, protocol (s), re-configurability, reassembles, reference, relationship, SONET, requirement(s), scrambling, security, signalling, simulation, SNIR, specification(s), uses, standard(s), technologies, types, translate, transports , vendors, wired-interface, wireless, capabilities, chipsets, serial, components, constraints, impairment(s), controllers
OSI, offloading	presentation layer	coding, creates, entities, functions, help, software implementation(s), independent, interpretation, protocol (s), provides, service(s), typically
making, OSI, NetBIOS, several, protocols, model	session layer	constructs, creates, ensures, establishes, functions, handles , implementation(s), interface, protocol (s), tokens
Backplane -, physical -	layer controller	family
cross-	layer design	paradigm, principles
upper -, presentation , link , network, transport , MAC	layer entity(ies)	makes up, is encapsulated
physical , higher-, link ,	layer frame(s)	formats, relay, PPP
MAC , application, link , network, physical , presentation , session , transport , OSI	layer function(s)	typically, via, provide, equivalent, including
Cross-, application, control, higher, link , network, packet, session , physical , second, sockets , upper -, control, presentation , MAC , signalling, transport , application-, internet, lower -, chosen	layer protocol (s)	engineering, results, designed, translation, interaction, analysis, packet, addresses, defined, based, localTalk, type, information, supported, configuration, PLP , PHY , turns, suites, transparency , specified, ASN, -specific, running, TCP, include, used, offer, PPP , hide
space charge	layer SCL	boundaries, around
application, higher-, IP, link , MAC , network, physical , sockets , transport	layer security	features, functions, using, EAP, WTLS, TLS, mechanism, SSH
lower -, upper -, physical	layer signal(s)	power, uses, relative, sees,

		<i>reconstruction, using, carrier</i>
<i>secure socket(s)</i>	<i>layer SSL</i>	<i>certificate, provides, connection, clients, protocols</i>
<i>network, transport</i>	<i>layer tree</i>	

10.5 Análisis detallado de formas generales.

Al igual que entre familias de palabras académicas se observan comportamientos que inducen a realizar un análisis más detallado para comprobar su naturaleza, entre familias de palabras generales se advierten casos que suscitan dudas.

Entre las mil familias de palabras más significativas de la ingeniería de telecomunicación se encuentran familias generales, registradas en el listado de West (1953), en las que más de la mitad de los miembros son valorados como específicos cuantitativamente. Asimismo, el examen realizado detecta palabras que formalmente son propias del lenguaje general y a las que estadísticamente otorga una puntuación acorde. Sin embargo, cualitativamente, el resultado es cuestionable. En consecuencia, *banking, bus, host, hub, internet, linear, mapping, memory, noise, programming, radio, escenario, shell, signal, tuning y windows*, se someterán al análisis detallado correspondiente, puesto que se contempla la posibilidad de que hayan sufrido un cambio o adaptación del significado general a un sentido especializado.

El procedimiento seguido para el análisis sigue el modelo de *bandwith* introduciendo una variación. En el apartado dedicado al estudio de los colocados, las fichas de palabras generales incluyen los colocados significativos de la forma bajo estudio en el corpus general Lacell, con la finalidad de averiguar si los usos realizados en los dos registros coinciden o son diferentes. En el Apéndice VIII, se hallan las fichas de todas las formas analizadas junto a los listados completos de los colocados, colocados significativos en el corpus técnico y en el general, líneas de concordancia y clusters. A continuación se muestran las fichas correspondientes al análisis de: *signal, noise, linear, programming, windows, mapping, shell, hub y bus*.

1. Familia de *signal*. Forma analizada: SIGNAL

1. FRECUENCIA.

KEYWORD	Frec. Telec	Frec. Lacell	Ratio	Término, Chung	KEYNESS	P valor
3. SIGNAL / West	7.022	641	41,6047867	NO	17.922,60	0
Miembros relacionados:	F Telec	F Gene	Ratio	Término	Keyness	P valor
SIGNALS	2.893	393	27,9573747	NO	6.852,30	0
SIGNALING	601	22	103,751001	ESPECIFICO	1.705,00	0
SIGNALLING / West	170	88	7,33680123	NO	243,2	0
SIGNALED	21	17	4,69149089	NO	21,6	0,000003
SIGNALTONEISE	5	0	inf/esp	inf/esp	15,7	0,000075
Familia técnica: No, estadística y formalmente.						

2. DISTRIBUCIÓN.

Distribución por áreas	Distribución Keyness	Keykeyword	Nº de textos	Porcentaje
9	0	SIGNAL	348	21,04%
<i>Dispersion plot</i>				
Area	Words	Hits	Per 1,000	
2 Ar. Comp	329.643	1.165	3.53	
4 Signal proc.	580.936	1.941	3.34	
081 Esp. Sign	867.208	2.208	2.55	
1 Electronics	722.823	1.054	1.46	
7 Systems	307.691	120	0.39	
5 Materials	101.241	38	0.38	
082 Esp. Tele	997.727	346	0.35	
3 Telematics	1.205.064	226	0.19	
6 Business	373.079	11	0.03	

3. COLOCACIÓN.

3.1 Colocados.

N	Collocates	TOTAL	LEFT	RIGHT	L5	L4	L3	L2	L1	*	R1	R2	R3	R4	R5
1	THE	6663	4267	2396	665	536	504	1345	1217	0	25	713	560	501	597
2	A	2348	1419	929	186	175	276	359	423	0	11	268	220	209	221
3	OF	2314	1648	666	209	231	496	532	180	0	82	87	83	219	195
4	AND	1752	796	956	172	185	143	189	107	0	229	270	127	150	180
5	TO	1739	816	923	186	198	223	162	47	0	277	146	178	156	166
6	IS	1643	434	1209	187	112	90	39	6	0	530	183	161	159	176
7	IN	1209	577	632	144	98	127	146	62	0	107	169	130	121	105
8	PROCESSING	802	64	738	12	16	19	16	1	0	689	11	21	10	7
9	FOR	701	409	292	52	81	101	128	47	0	39	87	63	48	55
10	THAT	534	276	258	54	55	84	62	21	0	58	55	55	49	41
11	AS	512	232	280	59	50	57	48	18	0	41	63	56	76	44
12	BE	504	139	365	62	50	20	6	1	0	3	132	93	80	57
13	THIS	480	226	254	42	40	36	36	72	0	3	102	48	50	51
14	AN	473	295	178	34	34	50	177	0	0	11	48	44	43	32
15	WITH	468	212	256	48	35	61	54	14	0	77	44	45	40	50
16	BY	447	174	273	40	52	59	18	5	0	45	58	72	49	49

17	ON	398	208	190	39	26	57	58	28	0	57	30	39	38	26
18	INPUT	388	299	89	22	15	17	15	230	0	7	4	26	23	29
19	AT	378	82	296	27	20	15	19	1	0	89	81	40	44	42
20	FROM	370	130	240	24	18	45	35	8	0	92	42	46	35	25
21	DIGITAL	352	295	57	17	13	11	8	246	0	2	8	22	16	9
22	OR	337	149	188	27	27	22	44	29	0	34	30	41	41	42
23	ARE	315	112	203	43	37	15	16	1	0	23	51	52	40	37
24	CAN	312	75	237	25	26	18	3	3	0	65	42	46	36	48
25	TIME	301	180	121	25	33	19	45	58	0	6	7	35	44	29
26	FREQUENCY	294	137	157	27	54	19	18	19	0	43	18	35	28	33
27	NOISE	288	82	206	22	16	18	15	11	0	2	119	30	30	25
28	IT	259	63	196	21	24	11	5	2	0	2	48	62	41	43
29	VALUE	252	152	100	20	47	70	13	2	0	18	6	14	39	23
30	OUTPUT	250	178	72	13	16	9	14	126	0	9	6	23	20	14
31	ASSIGNMENT	241	28	213	6	11	9	1	1	0	179	21	0	4	9
32	RECEIVED	221	167	54	5	5	2	9	146	0	14	11	13	7	9
33	WHICH	207	90	117	16	32	19	21	2	0	23	43	14	14	23
34	ONE	198	97	101	27	15	4	10	41	0	2	20	37	16	26
35	WHEN	190	119	71	9	15	43	48	4	0	9	22	16	9	15

3.2 Colocados significativos.

A) Colocados significativos en el corpus técnico.

Nº	Colocados	Z	T	MI
1	WATERMARK	29,86	2,75	3,56
2	VHDLAMS	29,24	2,60	3,66
3	SIGNAL'S	53,79	5,85	3,08

B) Colocados significativos en el corpus general.

Nº	Colocados	Z	T	MI
1	WELL	30,89	2,22	7,59
2	VEHICULAR	73,08	2,64	9,58
3	TL	57,37	3,15	8,37
4	SENDING	32,48	2,81	7,06
5	PROCESSING	31,32	3,69	6,17
6	FREQUENCY	16,37	2,40	5,54
7	ELECTRICAL	12,69	2,17	5,09
8	DIGITAL	41,17	3,96	6,75
9	CROSSING	17,61	2,59	5,53
10	BRADLEY	30,53	2,22	7,56
11	BATTERY	24,72	2,62	6,48

3.3 Definición del nodo en el diccionario.

1. Significado de *signal* en la *General Service List* (West, 1953):

Signal (noun):

1. Give a signal.
2. A railway signal, fog signal.
3. Signal tower, box, etc.

Signal (verb):

1. Signal to him to begin.
2. Flag-signalling.

2. Definición de *signal* en el *Advanced Learner's Dictionary* de Cambridge (2003):

Signal (noun):

1. (action) *An action, movement or sound which gives information, a message, a warning or an order.*
2. (showing) *Something which shows that something else exists or is likely to happen.*
3. (equipment) *Equipment, especially on the side of a railway or road, often with lights, which tells drivers whether they can go, must stop or should move more slowly.*
4. (wave) *A series of electrical or radio waves which are sent to a radio or television in order to produce a sound, picture or message.*

Signal (verb):

1. *To make a movement, sound, flash, etc. which gives information or tells people what to do.*
2. **Signal** (adj.): *(important) noticeable; not ordinary.*

3. Definición de *signal* en el diccionario especializado Webster:

Signal (noun), **Specialty Definition:**

In computing:

A synchronous language by Le Guernic et al of INRIA. A predefined message sent between two Unix processes or from the kernel to a process. Signals communicate the occurrence of unexpected external events such as the forced termination of a process by the user. Each signal has a unique number associated with it and each process has a signal handler set for each signal. Signals can be sent using the kill system call.

In electrical engineering:

A frequency or combination of frequencies transmitted according to a code over a circuit.

3.4 Colocados contiguos.

a) Colocado + <i>signal</i> .		
Total colocados seleccionados: 51	Frecuencia máxima: 246	Frecuencia mínima: 14
<ul style="list-style-type: none"> - Sustantivos, total 21: <i>input 230, output 126, layer 67, time 58, clock 46, data 41, error 40, control 39, speech 34, reference 23, voice 21, pilot 20, code 19, delay 19, domain 17, radio 17, wave 17, baseband 15, carrier 15, radar 15, host 14.</i> - Adjetivos, total 16: <i>digital 246, small 87, concurrent 71, optical 57, analog 43/analogue 21, original 43, conditional 37, video 36, electrical 35, audio 19, high 19, internal 15, weak 15, new 14, complex 14.</i> - Participios verbales, total 9: <i>recieved 146, mixed 128, transmitted 67, desired 64, modulated 42, based 35, selected 34, sampled 28, incoming 16.</i> - Formas abreviadas, total 5: <i>E 34, PN 26, RF 26, QPSK 16, GTLP 14.</i> 		
b) <i>Signal</i> + colocado.		
Total colocados seleccionados: 32	Frecuencia máxima: 179	Frecuencia mínima: 12
<ul style="list-style-type: none"> - Sustantivos, total 29: <i>assignment 179, strength 71, power 60, level 51/levels 24, values 48, integrity 47, processors 42/processor 31, quality 36, analysis 30, paths 27/path 19, bandwidth 26, transmission 26, propagation 23, flow 22, generator 19, model 18, source 18, transitions 17, degradation 16, characteristics 15, amplitude 13, conductor 13, constellation 13, energy 12, simulation 12, theory 12.</i> - Participios verbales, total 3: <i>processing 689, conditioning 16, sent 12.</i> 		

3.5 Identificación de colocados contiguos especializados en las líneas de concordancia.

ircuit (Fig. 2a) driven by a generalized modulated signal: (6). The equivalent circuit in the general distribution of frequency versus PSD for a modulated signal. 8.3 Timing measurements. Timing m by the equation: (1) where s(t) is the modulated signal, Ac is the amplitude of the carrier (assum RMATION APPROACH. Any analog modulated signal (AM, FM or PM) can be described by the lipping. These phenomena distort the modulated signal and usually lead to spectral regrowth as d therefore, attention must be paid to modulated signal constellation and related peak-to-average lting signal will be the spectrum of the modulated signal convolved with the spectrum of the pulse t Fig. 7 with that of the spectrum of the modulated signal depicted in Fig. 8. A summary of the con

is bi-phase modulated. The resultant **modulated signal** envelope is not constant, and the typical . The system is driven with a digitally **modulated signal** (GMSK). The different blocks in this loop : (2). Expression (3) implies that the **modulated signal** in the time domain $u(t)$ can be represented "is . Expression (3) implies that the **modulated signal** in the time domain can be represented by per second required to represent the **modulated signal** in a signal space. Thus, one interpretation ect sequence (DS): with DS, a binary **modulated signal** is 'directly' multiplied by a code. The code this example, the 930.004 MHz FSK-**modulated signal** is mixed with another signal at 930.002 rams. Another way to view a digitally **modulated signal** is with an eye diagram. Separate eye dia ch in Fig. 1) and can observe how the **modulated signal** looks both in time and frequency domain. Figure 5 shows the extra hierarchical **modulated signal** power needed in order to guarantee a BE er density) of the signal. The resulting **modulated signal** resembles a low-level noise signal that is is operation. After despreading a data **modulated signal** results, and after demodulation the origina ponds to the ± 3 dB bandwidth of the **modulated signal**. $R_{s\delta 1} \text{ p aP}$ corresponds to the theoretical f any general linear circuit driven by a **modulated signal** such as AM, FM or PM. Simulation time f any general linear circuit driven by a **modulated signal** such as amplitude modulation, frequency its of any R, L, C circuit driven by any **modulated signal**. The proposed equivalent circuits that wer ch insight into the quality of a digitally **modulated signal**. They can also pinpoint the causes for an ifferent Ways of Looking at a Digitally **Modulated Signal** Time and Frequency Domain View. 5.1 P ading factor, is then overlayed on the **modulated signal** to ensure constant signal bandwidth} and different ways of looking at a digitally **modulated signal**. To examine how transmitters turn on an onductance, fed by a current-source **modulated signal**, will be (real) (13) (imaginary) (14). Follow ductance G, fed by a current source **modulated signal**, will be: (13). Following this approach, S modeled with the multiplication of the **modulated signal** with a pulse train of 50 % duty cycle at th of the receiver to detect the desired **modulated signal** without exceeding a given degradation du

Concordancias de *modulated signal*.

a) Colocado + *signal*:

<i>analog(ue) signal</i>	<i>error signal</i>	<i>radio signal</i>
<i>audio signal</i>	<i>GTLF signal</i>	<i>received signal</i>
<i>baseband signal</i>	<i>high signal</i>	<i>reference signal</i>
<i>based signal</i>	<i>host signal</i>	<i>RF signal</i>
<i>carrier signal</i>	<i>input signal</i>	<i>sampled signal</i>
<i>clock signal</i>	<i>internal signal</i>	<i>selected signal</i>
<i>code signal</i>	<i>large signal</i>	<i>small signal</i>
<i>complex signal</i>	<i>layer signal</i>	<i>speech signal</i>
<i>concurrent signal</i>	<i>mixed signal</i>	<i>time signal</i>
<i>conditional signal</i>	<i>modulated signal</i>	<i>transmit signal</i>
<i>control signal</i>	<i>optical signal</i>	<i>transmitted signal</i>
<i>data signal(s)</i>	<i>output signal</i>	<i>video signal</i>
<i>delay signal</i>	<i>pilot signal</i>	<i>voice signal</i>
<i>digital signal</i>	<i>PN signal</i>	<i>wave signal</i>
<i>domain signal</i>	<i>QPSK signal</i>	<i>weak signal</i>
<i>electrical signal</i>	<i>radar signal</i>	

b) *Signal* + colocado:

<i>signal amplitude</i>	<i>signal flow(s)</i>	<i>signal quality</i>
<i>signal analysis</i>	<i>signal generator</i>	<i>signal simulation</i>
<i>signal assignment(s)</i>	<i>signal integrity</i>	<i>signal source</i>
<i>signal bandwidth</i>	<i>signal level(s)</i>	<i>signal strength</i>
<i>signal characteristics</i>	<i>signal model</i>	<i>signal theory</i>
<i>signal conditioning</i>	<i>signal path(s)</i>	<i>signal transitions</i>
<i>signal conductor(s)</i>	<i>signal power</i>	<i>signal transmission</i>
<i>signal constellation</i>	<i>signal processing</i>	<i>signal values</i>
<i>signal degradation</i>	<i>signal processor(s)</i>	
<i>signal energy</i>	<i>signal propagation</i>	

4. AGRUPACIONES LÉXICAS O CLUSTERS.

4.1 Distribución de clusters.

Combinaciones	Volumen	Frecuencia máxima
2-cluster	612	1250
3-cluster	1151	279
4-cluster	814	48
5-cluster	450	30
6-cluster	298	8

4.2 Inclusión de combinaciones especializadas en unidades más amplias.

<p><i>analog(ue) signal</i> <i>analog signal processing</i> <i>original analogue signal</i> based signal <i>based signal processing</i> clock signal <i>arbitrary clock signal</i> <i>rate clock signal</i> <i>edge of the clock signal</i> <i>generate an arbitrary clock signal</i> <i>logical function of the clock signal</i> concurrent signal <i>concurrent signal assignment(s)</i> <i>simple concurrent signal</i> <i>concurrent signal assignments statement(s)</i> <i>simple concurrent signal assignment</i> <i>simple concurrent signal assignment statements</i> <i>concurrent signal assignment statements conditional</i> domain signal <i>time domain signal</i> GTLT signal <i>GTLT signal levels</i> <i>interface between GTLT signal</i> <i>directional interface between GTLT signal</i> <i>GTLT signal levels and LVTTL</i> <i>interface between GTLT signal levels</i> <i>banckplane operating at GTLT signal</i> <i>GTLT signal and LVTTL logic</i> high signal <i>high signal transitions</i> input signal(s) <i>input signal values</i> <i>input signal changes</i> <i>input signal checked</i> <i>input signal frequency</i> <i>input signal level</i> <i>input signal select</i> <i>noise input signal</i> <i>different input signal</i> <i>input signal amplitude</i> <i>sinuisoidal input signal</i> <i>layer signal</i> <i>low layer signal</i> <i>layer signal power</i> pilot signal <i>pilot signal component</i> QPSK signal <i>legacy QPSK signal</i> radar signal <i>radar signal processing</i> <i>radar signal characteristics</i> <i>victim radar signal</i> radio signal</p>	<p>conditional signal <i>conditional signal assignment(s)</i> <i>conditional signal assignment statements</i> control signal <i>processing control signal</i> digital signal(s) <i>digital signal processing</i> <i>digital signal processor(s)</i> <i>digital signal level</i> <i>programmable digital signal</i> <i>digital signal integrity</i> <i>digital signal conversion</i> <i>DSP digital signal</i> <i>fast digital signal</i> <i>digital signal processing DSP</i> <i>digital signal processor(s) DSP(s)</i> <i>digital signal level (n°)</i> <i>fundamentals of digital signal</i> <i>digital signal processing techniques</i> <i>guide to digital signal</i> <i>software programmable digital signal</i> <i>advances in digital signal</i> <i>digital signal processing engineers</i> <i>digital signal processing systems</i> <i>programmable digital signal processors</i> <i>introduction to digital signal processing</i> <i>digital signal processing DSP applications</i> <i>fundamentals of digital signal processing</i> <i>engineer's guide to digital signal</i> <i>guide to digital signal processing</i> <i>applied introduction to digital signal</i> <i>data acquisition and digital signal</i> <i>engineer's guide to digital signal processing</i> <i>applied introduction to digital signal processing</i> <i>data acquisition and digital signal processing</i> <i>time data acquisition and digital signal</i> mixed signal <i>analogmixed signal</i> <i>analogmixed signal simulation</i> <i>mixed signal design</i> modulated signal <i>digitally modulated signal</i> <i>analog modulated signal</i> <i>generalized modulated signal</i> <i>modulated signal AM</i> <i>source modulated signal</i> <i>analog modulated signal AM</i> optical signal <i>optical signal power</i> <i>optical signal-to-noise</i> <i>optical signal-to-noise ratio</i> <i>electrical signal into an optical signal</i> <i>optical signal to noise ratio OSNR</i></p>
---	--

<p>radio signal strength received signal received signal strength received signal power mean received signal mean received signal strength RF signal radio frequency RF signal selected signal selected signal assignment selected signal assignment statement(s) conditional and selected signal small signal small signal characteristics small signal discrete small signal parameters small signal models small signal discrete device(s) signal processing digital signal processing time signal processing signal processing algorithms signal processing techniques speech signal processing signal processing methods signal processing research analog signal processing signal processing community signal processing applications IEEE signal processing receiver signal processing signal processing systems adaptive signal processing biomedical signal processing multirate signal processing statistical signal processing modern signal processing signal processing tools new signal processing advance signal processing radar signal processing signal processing operations signal processing perspectives signal processing tasks signal processing technology sophisticated signal processing array signal processing intensive signal processing signal processing chair signal processing engineer(s) signal processing paradigm signal processing problems signal processing required signal processing researchers signal processing technologies signal processing toolset traditional signal processing based signal processing field signal processing implementing signal processing multiresolution signal processing signal processing capability signal processing circuitry signal processing devices signal processing group signal processing laboratory signal processing literature</p>	<p>output signal output signal port output signal parameter output signal values output signal and/or one time signal time signal processing discrete time signal discrete time signal processing time signal processing paradigm time signal and sampling rate time signal and sampling rate conversion speech signal(s) speech signal processing time speech signal time speech signal processing speech signals and speech signal processing video signal analog video signal wave signal microwave signal sine wave signal square wave signal signal constellation(s) QAM signal constellation(s) signal flow(s) signal flowgraph signal flow graph(s) signal flow graph representation signal generator(s) signal generator blocks signal integrity improves signal integrity digital signal integrity signal integrity issues circuitry improves signal integrity OECTM circuitry improves signal integrity signal level(s) GTLF signal levels digital signal level signal level (n°) input signal level signal levels and LVTTL digital signal level (n°) signal path(s) long signal path widths of signal paths lead to long signal paths signal processor(s) digital signal processor(s) signal processor DSP companding signal processor ELIN signal processors linear signal processors digital signal processor(s) DSP(s) programmable digital signal processors signal propagation signal propagation mode signal propagation network signal source(s) weak signal source signal theory signal theory and communications signal theory and telecommunication signal transitions high signal transitions signal value(s)</p>
--	---

<p>signal processing means various signal processing VLSI signal processing digital signal processing DSP time speech signal processing discrete time signal processing signal processing DSP applications signal processing with fractals advanced receiver signal processing digital signal processing techniques time signal processing paradigm communications and signal processing computational intensive signal processing digital signal processing engineers digital signal processing systems domain of signal processing implementing signal processing algorithms intensive signal processing algorithms modern signal processing techniques receiver signal processing methods introduction to digital signal processing fundamentals of digital signal processing algorithms for array signal processing algorithms for statistical signal processing guide to digital signal processing computationally intensive signal processing algorithms analog integrated circuits and signal processing engineer's guide to digital signal processing squares algorithms for digital signal processing applied introduction to digital signal processing</p>	<p>input signal values store signal values output signal values value signal based necessary to store signal values flops necessary to store signal values signal amplitude input signal amplitude signal analysis small-signal analysis signal assignment(s) conditional signal assignment(s) selected signal assignment minimize signal assignment signal assignments statements simple signal assignment concurrent signal assignments conditional signal assignment statement(s) selected signal assignments concurrent signal assignment statement(s) selected signal assignment statements simple concurrent signal statement sequence of signal assignment branches of signal assignment minimize signal assignment statements signal power received signal power layer signal power optical signal power</p>
---	---

4.3 Colocados contiguos de las combinaciones especializadas.

PRE +	COMBINACIÓN	+ POST
equivalent, original, received, continuous, CMOS, CAS, mode, speed, varying, filtered, original, rapid	analog(ue) signal	comprising, does, passes, processing, reconstruction, output, transitions
selected, digital, small, live	audio signal	components, digitized, enhancement, management, processing, sampled, strength, transmission
original, terminal, received, desired	baseband signal	model, processing, upper
knowledge-, sample-, frame-, wavlet	based signal	analysis, processing
single-, high-frequency, RF, amplitude	carrier signal	arriving, lower, moves
separate, bit-rate, arbitrary, global, Hz	clock signal	arrival, computation, determines, local, shown
generated, PN, rate, discrete-valued, SD	code signal	assigned, consists, due to, time
incomplete	conditional signal	assignment(s), energy, statement
processing, corresponding, hazard, electric, saturated, gate, powerful, processing, MemRead, provides, defect, appropriate, defect, computed	control signal	applied, becomes, Ci, CLK, driving, exceeds, indicating, routing, sets
modulated, NRZ, binary, intended, despread, transmitted, multiplexed, unfiltered	data signal(s)	DataProg, modulates, occupies, respectively, input, transmitted
transformed, original, estimated, time-varying, online, time-, correlated	delay signal	
modulating, fast, Q, analog-to-, PN, DSI, operating-rate, channelized, high-performance, custom, Java, new, integrated, designing, practicable,	digital signal	according, cables, clean-up, clocks, compression, conditioning, conversion, integrity, level, processing, processors, provided, pulse, transmission(,)

<i>practical, transform-based, advanced, find, build, using, performing, programmable, recent, real-VLSI, pulse, next-generation, time, DSP, two-dimensional</i>		<i>transport, connectivity, phase, small</i>
<i>time, frequency-, time-, cross-,</i>	<i>domain signal</i>	<i>envelope, integrity</i>
<i>continuous, optical-to-, E1/T1, E1, uses, same, information-bearing, send, equivalent, analog</i>	<i>electrical signal</i>	<i>carried, conversion, levels, propagation, sent, voltage</i>
<i>motion, desired, global, standard, residual, filtered, resulting, bus, mean</i>	<i>error signal</i>	<i>analysis, computed, derived, driving, slowly, takes, um</i>
	<i>GTLP signal</i>	<i>levels</i>
<i>maintaining, having, low-to-</i>	<i>high signal</i>	<i>frequencies, gains, integrity, levels, power, swings, -to-noise, transitions</i>
<i>Gaussian</i>	<i>host signal</i>	<i>statistics</i>
<i>DC, noise, control, sinusoidal, modulation, traces, zero, periodic, noise, SDI, strong, carrier, differential, composite, load, large, small, analog, in-phase, opcode, single-, wave, peak-to-peak, complex, state, wideband, ended, electrical, channel, noisy, current, filter, sampled, valid</i>	<i>input signal</i>	<i>adaptation, amplitude, bandwidth, changes, checked, clk, contains, cycle, D, described, drives, experiences, frequency, helps, level(s), lies, magnitude, named, occurs, -output, provided, sees, select, sources, spectral, statistics, strength, time, VI, values, vector, waveform</i>
	<i>internal signal</i>	<i>lines</i>
<i>exercise, relate, MOS</i>	<i>large signal</i>	<i>behaviour, behaviour, input, nonidealities, operation, performance, quantities, response, sense, steady-state, swings, amplifiers, base, circuits, diode, integrator, limitation, model, nonlinearities, response, simulations, steady, transient</i>
<i>upper-, lower-</i>	<i>layer signal</i>	<i>carrier, corrupted, power, processor, reconstruction, relative, sees, shown, space, uses, using, transmission</i>
<i>supplying, develop, AMCC's, international, IEEE, selecting, analog, performance, complex, facilitate, concurrent, analog-, innovative, RF, future, analog-intensive, today's, evaluating, accurate, international, full, general, include</i>	<i>mixed signal</i>	<i>analogue, chip(s), circuit(s), CMOS, design(s), devices, fuzzy, HDLs, layout, processors, product, programmable, specification, testing, VLSI, Benchmark, CAS, capabilities, components, designers, devices, domains, equation, extensions, fault, functional, IC, integrated, implementations, interfacing, languages, modelling, MS, MOS, nature, products, RF, simulation(s), SoC, system(s), technology, testing, types, VHDL-AMS</i>
<i>generalized, analog, digitally, resultant, binary, FSK-, hierarchical, resulting, data, source, current-source, desired 10G, reduce, transmitted, modulated, received,</i>	<i>modulated signal</i>	<i>AM, constellation, convolved, depicted, envelope, GMSK, time, looks, power, resembles, results,</i>
<i>output, power, resulting, transmitted, incoming, all-, dispersed, detectable</i>	<i>optical signal</i>	<i>bandwidth, carrier, integrity, power, processing, see, -to-noise, transmission</i>
<i>DAC, lower, source-decoder, desired, sum, plant, model, linear, attenuated, complex, composite, calculates, required, filter, compute, integrated, short</i>	<i>output signal</i>	<i>accordingly, changes, decays, depends, distorted, does, graph, GVG, level, menu, parameter, paths, phase, port, Q, sent, setting, swing, takes, value(s), Vout, Vd</i>
	<i>pilot signal</i>	<i>component, dedicated, does</i>
<i>interfering, exact, different, same, shifted, basestation's, else's, looping, same</i>	<i>PN signal</i>	<i>middle, results, shifted</i>

<i>layer, upper-layer, legacy, high-priority, narrowband, keying</i>	<i>QPSK signal</i>	<i>becomes</i>
<i>all, victim, police, threat, transmitted</i>	<i>radar signal</i>	<i>characteristics, gets, processing</i>
<i>microwave, via, unique, mobile, integral, RSSI, remote</i>	<i>radio signal</i>	<i>processor, received, strength</i>
<i>joint, first, desired, proper, multi-satellite, allowable, actual, target, user's, poor, mean, average, zero</i>	<i>received signal</i>	<i>collected, consists, detection, drops, due to, envelope, frequency, given, input, level(s), power(s), quality, resolves, samples, sequence, strength, suffers, variance, vector, waveforms</i>
<i>known, output, external, ideal, accurate, offset</i>	<i>reference signal</i>	<i>changes, occurs, producing</i>
<i>instruments', modulated, compressed, frequency, quality, input, incoming</i>	<i>RF signal</i>	<i>affects, analyzer, coming, generation, interface</i>
<i>original, periodically, simulated, resulting</i>	<i>sampled signal</i>	<i>obtained, prior</i>
<i>concurrent, first</i>	<i>selected signal</i>	<i>assignment(s)</i>
<i>AC, BJT, term, expected, same, measured, using, identical, slope, low-frequency, model, obtain, classical, maximum</i>	<i>small signal</i>	<i>absorption, analysis, behaviour, behavioural, characteristics, discrete(s), drain-source, excursions, frequency, gain, input, model, MOSFETs, operation, output, parameters, processing, regime, resistance, sensitivity, silicon, simulation, transfer, values, AC, analysis, approximation, circuit, class, equivalent, model, parameters, properties, simulation</i>
<i>encoded, analog, discrete-time, digital, encoded</i>	<i>speech signal</i>	<i>processing, researchers, useful</i>
<i>continuous-, discrete-, real-, space-, height-</i>	<i>time signal</i>	<i>contains, discussed, generation, output, processing</i>
<i>GHz, far-end</i>	<i>transmit signal</i>	<i>entering, information, path, power, pulses, strength, cable</i>
<i>known, same, effective, original, near, B-MAC</i>	<i>transmitted signal</i>	<i>amplitude, arrive, bandwidth, energy, flows, illuminates, quality, reflects, results, varies</i>
<i>analog, downstream, conventional, couple, broadcast, colour, aggregated, bit, develop, selected</i>	<i>video signal</i>	<i>digitization, enhancement, processing, processors, via</i>
<i>Original, digital, CDMA</i>	<i>voice signal</i>	<i>processing, spread</i>
<i>square, sine</i>	<i>wave signal</i>	<i>lying, select</i>
<i>league, hydrogen, line, calibrated, MHz</i>	<i>weak signal</i>	<i>source, amplification</i>
<i>basic, input, message, Rayleigh, represents, peak, transmitted</i>	<i>signal amplitude</i>	<i>changes, given, scaling, varies, scaling</i>
<i>cyclostationary, error, nonstationary, small, biomedical, extraction, knowledge-based, transient, time-frequency, spectral</i>	<i>signal analysis</i>	<i>based, methods</i>
<i>concurrent, conditional, following, minimize, multiple, selected, simple, successive, uses, using, VHDL, concurrent, end, example, formulate, sequential</i>	<i>signal assignment(s)</i>	<i>statement(s), constructing, selected, conditional</i>
<i>higher, input, optical, original, ST, total, reduced, desired, transmitted, constant</i>	<i>signal bandwidth</i>	<i>capability, given</i>
<i>radar, small, nominal, nonstationary</i>	<i>signal characteristics</i>	<i>including, determined, change</i>
<i>card's, front-end, specifications, complex, high-speed, digital</i>	<i>signal conditioning</i>	<i>system, equipment, products, radar, circuitry, components, accessory</i>
<i>coax, equivalent, breaking, N, P</i>	<i>signal conductor(s)</i>	<i>width, discontinuity, chosen, contain</i>
<i>QAM, higher, larger, modulated</i>	<i>signal constellation</i>	<i>according, displays, uses</i>
<i>acceptable, overall</i>	<i>signal degradation</i>	<i>due to, introduces</i>

<i>conditional, removes, transmitted</i>	<i>signal energy</i>	
<i>systems, using</i>	<i>signal flow(s)</i>	<i>graph(s), descriptions, diagram</i>
<i>today's, discuss, kHz, QAM, Simulink's, using, Hz, vector</i>	<i>signal generator</i>	<i>specifications, fully, block(s), need</i>
<i>large-, analyse, analyzing, autonomous, digital, maintain, cross-domain, exceptional, extensive, frequency, fundamental, improve(s), in-module, optical, specifically, superior, timing/, decreasing, performance, brings, high,</i>	<i>signal integrity</i>	<i>means, effects, early, issue(s), problems, monitoring, matters, features, management, limits, analysis, concern, closure</i>
<i>CDMA, decision, desired, digital, highest, input, lower, GTLP-to-LVITL, measured, output, GTLP-to-LVrrL, receive, GTLP-to-LVTrL, strongest, IF, received, upper, average, monitor, full-swing, transport, constant, boost, GTLP</i>	<i>signal level(s)</i>	<i>higher, ADSL, zero, applications, translation, needed, power, foes, drops, obtained, received</i>
<i>baseband, good, MLS, small, large, GPS</i>	<i>signal model(s)</i>	<i>propagation</i>
<i>cause, critical, long, transmit, connecting, simple, entire, dedicated, multiple, skew, straight, switched, wider</i>	<i>signal path(s)</i>	<i>design, known, linearity, distortion, consists, resulting</i>
<i>combined, desired, extra, high, initial, upper-layer, lower-layer, LOS, maximum, modulated, net, received, optical, ST, transmit, undesired, RMS, useful</i>	<i>Signal power</i>	<i>Pd, level(s), relative, needed, required, gains, decays</i>
<i>academic, adaptive, advanced, allow, analog, analogue, applications, applying, approximates, audio, auditory, automotive, available, baseband, biomedical, bit, certain, channel, combines, common, complex, communications, complexity, configurable, contemporary, conventional, cubic, current, cyclostationary, dedicated, desired, develop, digital, different, elaborate, essential, employ, exploit, extra, fast, flexible, FPGA, fundamental, Gbps, homomorphic, IEEE, implementing, important, improved, include, integrated, intensive, interception, introductory, key, largely, latest, making, modern, multirate, multiresolution, new, nonlinear, NI's, novel, objects, onboard, all-optical, outboard, pertinent, powerful, radar, recast, receiver, related, represent, scale, significant, small, sophisticated, eigen-spaced, special, specialized, specifying, speech, standard, statistical, superior, real-time, discrete-time, traditional, transmitter, useful, various, video, VLSI, wavelet, world, distributed, photronic, wave-field, non-Gaussian, geological, multimedia, array high-speed, , biomedical, coherent, communication, OFDM, smart, statistical</i>	<i>signal processing</i>	<i>operation, elements, textbook, laboratory, architecture, work, tool, researchers, applications, methods, algorithms, techniques, systems, research, technical, devices, circuits, domain, needs, circuitry, communities, disciplines, worksystem, tasks, electronics, institute, capacity, necessary, equipments, analyses, required, continues, employs, methodologies, scenarios, operations, laboratory, perspectives, hardware, course, chair, approaches, offers, texts, DSP, engineers, edition, tasks, dates, software, available, capability, design, chip, devices, implementation, concepts, required, using, specialists, involving, architects, society, magazine, workshop, classes, functions, issues, real, algorithm(s), conferences, group, challenges, literature, means, toolset, fundamentals, contribute, resulted, tend, textbook, ICICS, focuses, information, stage, community, library, accomplished, covers, seems, framework, provides, paradigm, describes</i>
<i>actual, companding, digital, ELIN, lower-layer, radio, integrated, linear, mixed, practicing, specific, statistical, video, larger, multimedia, high-order, wireless, particular</i>	<i>signal processor(s)</i>	<i>DSP, chips, users, design, demodulates, technology, means, DSPs, differ</i>
<i>electromagnetic, different, electrical, mode, multipath, overcome</i>	<i>signal propagation</i>	<i>velocity, network(s), parts, discontinuities, characteristics, delay(s), mode effects, paths, time, conductors</i>

<i>best, improve, maintain, monitor, numerous, poor, upstream, overall, transmitted, acceptable</i>	<i>signal quality</i>	<i>information, parameters, improvement, monitoring, based, decreasing, bit, measured</i>
<i>mixed-, modern, small-, full, large-</i>	<i>signal simulation</i>	<i>algorithms, techniques, SN-model</i>
<i>block, test, weak, input, separate, several</i>	<i>signal source(s)</i>	<i>kit, assigned</i>
<i>emplaced, greater, higher, largest, lower, radio, received, smaller, Jammer-to-, transmit, adequate, AP, detectable, display, input, highest, monitors, pattern, station's, audio, corresponding, different</i>	<i>signal strength</i>	<i>drops, target, index, dB, indicator(s), indication, check, occurring, varies, shows, information, deviation, measurements, increased</i>
<i>analogue, low-to-high, using</i>	<i>signal transitions</i>	<i>shown, provide</i>
<i>advanced, audio-, boost, different, digitizing, layered, novel, optimized, uniform, distance, optical, digital</i>	<i>signal transmission</i>	<i>voltage, procedure, improves, techniques, speeds, structure</i>
<i>correct, final, output, recent, associated, corresponding, example, input, new, small, undefined, uninitialized, write, implicit, incorrect, resolving, store, updating</i>	<i>signal value(s)</i>	<i>propagating, represents, unaffected, according, independent, change, based</i>

2. Familia de *noise*. Forma analizada: NOISE

1. FRECUENCIA.

KEYWORD	Frec. Telec	Frec. Lacell	Ratio	Término, Chung	KEYNESS	P valor
121. NOISE / West	1.893	878	8,18835385	NO	2.887,20	0
Miembros relacionados:	F Telec	F Gene	Ratio	Término	Keyness	P valor
DENOISING	16	0	inf/esp	inf/esp	50,2	0
NOISY / West	97	198	1,86057443	NO	23,1	0,000002
NOISING	7	1	26,585115	NO	16,4	0,000051
NOISEIT	5	0	inf/esp	inf/esp	15,7	0,000075
Familia técnica: No, estadística y formalmente.						

2. DISTRIBUCIÓN.

Distribución por áreas	Distribución Keyness	Keykeyword	Nº de textos	Porcentaje
9	0	NOISE	151	9,13%
<i>Dispersion plot</i>				
Area	Words	Hits	Per 1,000	
4 Signal proc.	580.936	497	0.86	
081 Esp. Sign	867.208	660	0.76	
1 Electronics	722.823	458	0.63	
7 Systems	307.691	80	0.26	
2 Ar. Comp	329.643	42	0.13	
5 Materials	101.241	11	0.11	
082 Esp.Tele	997.727	100	0.10	
6 Business	373.079	11	0.03	
3 Telematics	1.205.064	32	0.03	

3. COLOCACIÓN.

3.1 Colocados.

N	Collocates	TOTAL	LEFT	RIGHT	L5	L4	L3	L2	L1	*	R1	R2	R3	R4	R5
1	THE	1444	836	608	142	145	161	164	224	0	4	183	154	130	137
2	AND	608	270	338	36	52	52	78	52	0	105	82	55	54	42
3	OF	582	372	210	61	48	55	137	71	0	26	45	21	51	67
4	TO	544	365	179	47	51	57	64	146	0	16	35	39	49	40
5	IS	450	161	289	60	47	28	17	9	0	91	62	41	58	37
6	IN	374	122	252	38	34	25	19	6	0	79	60	38	38	37
7	A	361	176	185	48	30	34	43	21	0	0	42	49	44	50
8	SIGNAL	287	209	78	20	35	25	126	3	0	10	14	15	17	22
9	AS	171	87	84	14	18	15	26	14	0	12	23	20	16	13
10	WITH	158	100	58	16	22	21	31	10	0	10	15	11	13	9
11	FOR	152	86	66	21	18	15	24	8	0	2	24	17	10	13
12	LOW	137	117	20	2	11	14	15	75	0	0	5	6	5	4
13	THIS	137	58	79	14	13	9	4	18	0	0	30	18	14	17
14	BY	125	74	51	8	19	15	23	9	0	1	11	13	13	13
15	RATIO	121	8	113	4	0	3	1	0	0	92	7	10	1	3
16	BE	118	30	88	13	9	6	1	1	0	0	21	31	22	14
17	THAT	116	62	54	17	14	13	15	3	0	11	11	11	9	12

18	ARE	114	45	69	9	22	11	2	1	0	17	11	18	9	14
19	POWER	103	27	76	5	6	9	6	1	0	41	5	11	12	7
20	FROM	98	42	56	5	7	5	17	8	0	18	13	11	7	7
21	WHITE	92	82	10	1	4	1	34	42	0	0	2	4	2	2
22	AT	87	21	66	10	3	5	3	0	0	14	21	17	7	7
23	GAUSSIAN	87	73	14	1	0	0	4	68	0	0	3	2	2	7
24	HIGH	85	56	29	14	9	16	8	9	0	0	2	10	9	8
25	OR	84	38	46	2	7	10	13	6	0	12	15	6	7	6
26	INTERFERENCE	83	30	53	5	8	7	7	3	0	7	31	4	7	4
27	CAN	79	15	64	7	4	2	2	0	0	17	10	13	13	11
28	AN	74	36	38	4	17	4	11	0	0	0	11	14	4	9
29	ON	74	29	45	13	4	5	6	1	0	15	8	4	12	6
30	ADDITIVE	71	65	6	1	2	27	10	25	0	0	1	2	2	1
31	PHASE	67	57	10	2	0	2	0	53	0	2	1	5	1	1
32	FIGURE	65	7	58	1	3	2	0	1	0	44	4	2	5	3
33	THERMAL	62	52	10	1	1	5	0	45	0	0	1	3	1	5
34	IT	61	26	35	9	11	3	3	0	0	0	9	12	7	7
35	FREQUENCY	60	41	19	7	12	4	6	12	0	0	2	5	7	5

3.2 Colocados significativos.

A) Colocados significativos en el corpus técnico.

Nº	Colocados	Z	T	MI	Nº	Colocados	Z	T	MI
1	SQNR	38,10	2,23	8,19	55	CORRELATED	13,87	3,01	4,40
2	FLICKER	72,29	4,23	8,19	56	MULTIPLIED	10,66	2,33	4,38
3	IMPULSIVE	85,66	5,27	8,04	57	SHAPED	12,01	2,69	4,32
4	ADDITIVE	111,60	8,38	7,47	58	SHAPING	13,20	3,00	4,27
5	SHOT	67,04	5,35	7,29	59	OSCILLATORS	10,07	2,32	4,24
6	COLORED	39,08	3,96	6,61	60	RANDOM	30,83	7,20	4,20
7	SNIR	21,14	2,21	6,51	61	HARMONIC	10,69	2,50	4,19
8	SNR	61,67	6,48	6,50	62	STATIONARY	13,15	3,13	4,14
9	IMMUNITY	29,43	3,13	6,47	63	RESIDUAL	10,38	2,49	4,12
10	WHITE	88,46	9,48	6,44	64	MAI	9,51	2,31	4,09
11	AWGN	45,07	4,84	6,44	65	PN	13,74	3,39	4,04
12	MARGINS	36,12	3,95	6,38	66	DB	24,82	6,15	4,02
13	FOLDED	19,90	2,21	6,34	67	SPECTRAL	20,64	5,21	3,97
14	CANCELING	21,60	2,42	6,32	68	AMPLIFICATION	9,42	2,46	3,87
15	APD	26,05	2,96	6,27	69	MEASUREMENT	19,73	5,40	3,74
16	JAMMING	41,13	4,73	6,24	70	TONE	8,22	2,26	3,72
17	PASSIVES	22,31	2,61	6,19	71	TENDS	8,19	2,26	3,71
18	AMBIENT	36,51	4,30	6,17	72	DEGRADATION	11,21	3,18	3,63
19	CORRUPTED	30,71	3,69	6,12	73	PLUS	12,31	3,55	3,59
20	OVERSAMPLING	19,82	2,41	6,08	74	CONTRIBUTION	10,01	2,90	3,58
21	GAUSSIAN	73,49	9,18	6,00	75	AMPLIFIERS	11,84	3,43	3,58
22	COMFORT	20,50	2,60	5,95	76	WIDEBAND	8,22	2,42	3,53
23	JOHNSON	31,57	4,06	5,92	77	AMPLIFIER	17,62	5,24	3,50
24	MARGIN	41,06	5,30	5,91	78	OSCILLATOR	9,68	2,88	3,50
25	SUSCEPTIBLE	22,62	3,10	5,73	79	PHENOMENA	6,80	2,04	3,48
26	QUANTISATION	15,56	2,19	5,66	80	GENERATOR	10,38	3,15	3,45
27	RATIO	73,44	10,76	5,54	81	DETERMINISTIC	7,93	2,40	3,44
28	SPURS	14,55	2,19	5,47	82	SPURIOUS	7,29	2,22	3,43
29	LNA	21,36	3,24	5,44	83	DENSITY	15,42	4,71	3,42
30	THERMAL	50,03	7,69	5,40	84	NARROWBAND	6,59	2,02	3,40
31	QUANTIZATION	24,88	3,90	5,35	85	SENSITIVITY	11,38	3,51	3,40

32	CROSSTALK	21,07	3,38	5,28	86	SENSITIVE	11,27	3,50	3,37
33	RATIOS	23,50	3,77	5,28	87	PRESENCE	12,26	3,91	3,29
34	SUPPRESS	13,32	2,18	5,23	88	INDUCED	6,79	2,19	3,26
35	UNCORRELATED	13,15	2,17	5,19	89	FIGURES	10,35	3,35	3,25
36	RADIOMETER	12,81	2,17	5,12	90	REDUCING	10,63	3,46	3,24
37	DISTORTION	31,54	5,40	5,09	91	CONTRIBUTIONS	7,26	2,37	3,24
38	SKY	18,73	3,22	5,08	92	MULTIPATH	9,04	2,96	3,22
39	HEADSET	12,35	2,17	5,02	93	DISTINGUISH	7,14	2,36	3,20
40	VARIANCE	26,34	4,83	4,89	94	FILTERING	11,43	3,78	3,19
41	PSEUDO	19,69	3,62	4,89	95	ESTIMATION	12,91	4,27	3,19
42	PLL	13,92	2,56	4,89	96	AMPLITUDE	10,59	3,55	3,15
43	FLOOR	22,09	4,10	4,86	97	IMPULSE	8,37	2,81	3,15
44	BACKGROUND	30,31	5,93	4,71	98	HP	6,45	2,17	3,14
45	ENHANCEMENT	16,03	3,19	4,66	99	OBSERVATION	6,93	2,34	3,13
46	FLUCTUATIONS	11,59	2,35	4,61	100	REDUCTION	11,58	3,95	3,10
47	LINEARITY	14,58	3,03	4,54	101	CONSIDERATIONS	7,28	2,50	3,09
48	MIXER	15,20	3,17	4,52	102	INHERENT	6,73	2,33	3,06
49	INTERFERENCE	41,31	8,71	4,49	103	ADEQUATE	6,71	2,33	3,06
50	REJECTION	12,82	2,70	4,49	104	SINE	6,68	2,33	3,04
51	CANCELLATION	15,70	3,31	4,49	105	HZ	8,73	3,04	3,04
52	RIN	10,05	2,14	4,47	106	OFFSET	7,93	2,78	3,03
53	LTI	10,05	2,14	4,47	107	STABILITY	8,22	2,90	3,00
54	EXCESS	17,26	3,70	4,45					

B) Colocados significativos en el corpus general.

Nº	Colocados	Z	T	MI
1	DEAFENING	54,40	2,44	8,95
2	WOO	49,15	2,82	8,25
3	LOUDER	52,39	3,59	7,74
4	VIBRATION	35,06	2,81	7,28
5	BORNE	42,77	3,44	7,27
6	PITCHED	32,14	2,63	7,23
7	LOUD	59,73	5,25	7,02
8	INDUCED	19,59	2,21	6,30
9	RATS	18,32	2,41	5,86
10	EXCESSIVE	20,45	2,78	5,76
11	POLLUTION	17,81	2,76	5,38
12	SOUNDED	12,91	2,17	5,14
13	EXPOSED	13,90	2,56	4,89
14	FREQUENCY	11,49	2,16	4,83
15	TALE	11,05	2,15	4,72
16	TREMENDOUS	11,92	2,35	4,68
17	LIGHTING	9,83	2,13	4,41
18	DUST	10,55	2,33	4,36
19	AIRCRAFT	9,40	2,12	4,29
20	HELLO	8,42	2,10	4,01

3.3 Definición del nodo en el diccionario.

1. Significado de *noise* en la *General Service List* (West, 1953):

Noise (noun): *Make a noise. The noise of trains on the bridge. Noise of a banging door.*

2. Definición de *Noise* en el *Advanced Learner's Dictionary* de Cambridge (2003):

Noise (noun):

1. *Sound, especially when it is unwanted, unpleasant or loud.*

2. (Signal) (specialized): Any unwanted change in a signal, especially in a signal produced by an electronic device: Using a single chip reduces (the) noise on the output signal by 90%.

3. Definición de *noise* en el diccionario especializado Webster:

Noise (noun):

1. Sound of any kind (especially unintelligible or dissonant sound); "he enjoyed the street noises"; "they heard indistinct noises of people talking".
2. The auditory experience of sound that lacks musical quality; sound that is a disagreeable auditory experience; "modern music is just noise to me".
3. Electrical or acoustic activity that can disturb communication.

Specialty Definition:

In aerospace:

1. Any undesired sound. By extension, noise is any unwanted disturbance within a useful frequency band, such as undesired electric waves in a transmission channel or device. When caused by natural electrical discharges in the atmosphere, noise may be called static.
2. An erratic, intermittent, or statistically random oscillation.
3. In electrical circuit analysis, that portion of the unwanted signal which is statistically random, as distinguished from hum, which is an unwanted signal occurring at multiples of the power-supply frequency. If ambiguity exists as to the nature of the noise, a phrase such as acoustic noise or electric noise should be used. Since the above definition are not mutually exclusive, it is usually necessary to depend on context for the distinction.

In electrical engineering:

Any undesired signal, by extension any unwanted disturbance within a useful frequency band; a disturbance that affects a signal and that may distort the information carried by the signal.

3.4 Colocados contiguos.

a) Colocado + <i>noise</i> .		
Total colocados seleccionados: 26	Frecuencia máxima: 53	Frecuencia mínima: 6
<ul style="list-style-type: none"> - Sustantivos, total 17: <i>phase 53, random 44, background 30, shot 20, measurement 19, ambient 12, band 11, flicker 11, quantization 11, sky 10, excess 9, Johnson 8, impulse 7, radar 7, comfort 6, crosstalk 6, output 6.</i> - Adjetivos, total 7: <i>Gaussian 68, thermal 45, white 42, impulsive 25, better 8, external 7, total 6.</i> - Participios verbales que actúan como adjetivos, total 2: <i>colored 8, correlated 6.</i> - Prefijos, total 1: <i>pseudo 10.</i> - Formas abreviadas, total 1: <i>F 11.</i> 		
b) <i>noise</i> + colocado.		
Total colocados seleccionados: 36	Frecuencia máxima: 92	Frecuencia mínima: 6
<ul style="list-style-type: none"> - Sustantivos, total 30: <i>ratio 92/ratios 12, figure 44/figures 11, power 41, level 29/levels 15, margin 23/margins 16, floor 15, analysis 13, variance 13, process 12, reduction 12, amplifier 11, performance 11, sources 11/source 10, characteristics 9, components 9, current 9, contribution 8, immunity 8, measurements 8, sequence 8, voltage 8, enhancement 7, cancellation 6, generator 6, signals 6.</i> - Adjetivos, total 1: <i>spectral 6.</i> - Participios verbales, total 4: <i>jamming 10, shaping 10, canceling 6, filtering 6.</i> - Formas abreviadas, total 1: <i>AWGN 14.</i> 		

3.5 Identificación de colocados contiguos especializados en las líneas de concordancia.

*r to establish a practical specification for a low **noise amplifier** and down converter. In the previous results are applied to the case of a low **noise amplifier** (LNA). Particularly, it is shown , leads to smaller noise figure for the final low **noise amplifier** (LNA). Moreover, the pair Zloaches assume high quality passives. Thus, low **noise amplifier** designs primarily seek the traistors operating in efficient class 'C' mode. Low **noise amplifier** (LNA), digital attenuator and sh with High Performance and Low Cost. The low **noise amplifier** and mixer represent one of the Typically these include the antenna to the low-**noise amplifier** (LNA), power-amplifier output (*

- (-106 + 8.7 - 10) dBm = + 35.3dBm (8). *Low-Noise Amplifier and Downconverter with High*
oint of the engineer who wants to design a low-noise amplifier with either discrete transistors
ults are presented for MCM filters and for a low-noise amplifier block designed for the HIPERL
very next thing it does is get amplified by a low noise amplifier. Then the downconverter lowers
timal. In effect, as presented in this paper, low noise amplifiers are optimally designed if back
MIC and MMIC Cryogenic LNA's. Low noise amplifiers using InP-HEMT can be built
ded ready for circuit board integration with low noise amplifiers. Rectangular patch antennas
, including development of photoparametric low-noise amplifiers/converters for optical receivers,

Concordancias de noise amplifier(s).

a) Colocado + noise:

<i>ambient noise</i>	<i>Gaussian noise</i>	<i>quantization noise</i>
<i>background noise</i>	<i>impulse noise</i>	<i>radar noise</i>
<i>band noise</i>	<i>impulsive noise</i>	<i>random noise</i>
<i>colored noise</i>	<i>Johnson noise</i>	<i>shot noise</i>
<i>comfort noise</i>	<i>measurement noise</i>	<i>sky noise</i>
<i>correlated noise</i>	<i>output noise</i>	<i>thermal noise</i>
<i>excess noise</i>	<i>phase noise</i>	<i>total noise</i>
<i>external noise</i>	<i>pseudo noise</i>	<i>white noise</i>
<i>F noise</i>		
<i>flicker noise</i>		

b) Noise + colocado:

<i>noise amplifier</i>	<i>noise filtering</i>	<i>noise process</i>
<i>noise analysis</i>	<i>noise floor</i>	<i>noise ratio(s)</i>
<i>noise AWNG</i>	<i>noise free</i>	<i>noise reduction</i>
<i>noise canceling</i>	<i>noise generator</i>	<i>noise sequence</i>
<i>noise cancellation</i>	<i>noise immunity</i>	<i>noise shaping</i>
<i>noise characteristics</i>	<i>noise jamming</i>	<i>noise signals</i>
<i>noise components</i>	<i>noise level(s)</i>	<i>noise source(s)</i>
<i>noise contribution</i>	<i>noise margin(s)</i>	<i>noise spectral</i>
<i>noise current</i>	<i>noise measurements</i>	<i>noise temperature</i>
<i>noise enhancement</i>	<i>noise performance</i>	<i>noise variance</i>
<i>noise figure(s)</i>	<i>noise power</i>	<i>noise voltage</i>

4. AGRUPACIONES LÉXICAS O CLUSTERS.

4.1 Distribución de clusters.

Combinaciones	Volumen	Frecuencia máxima
2-cluster	218	238
3-cluster	234	9
4-cluster	95	27
5-cluster	35	14

4.2 Inclusión de combinaciones especializadas en unidades más amplias.

<p><i>phase noise</i> <i>low phase noise</i> <i>phase noise floor</i> <i>random noise</i> <i>add random noise</i> <i>added random noise</i> <i>random noise checkbox</i> <i>add random noise checkbox</i> <i>background noise</i> <i>background noise power</i> <i>shot noise</i> <i>shot noise and flicker noise</i> <i>shot noise and thermal noise</i> <i>thermal noise</i></p>	<p><i>sky noise</i> <i>sky noise temperature</i> <i>radar noise</i> <i>radar noise jamming</i> <i>Gaussian noise</i> <i>additive Gaussian noise</i> <i>coloured Gaussian noise</i> <i>Gaussian noise sequence</i> <i>additive white Gaussian noise</i> <i>white Gaussian noise AWGN</i> <i>Gaussian noise AWGN channel</i> <i>additive white Gaussian noise AWGN</i> <i>white Gaussian noise AWGN channel</i> <i>noise figure</i></p>
---	--

<p><i>thermal noise power</i> <i>thermal noise ratio</i> <i>thermal noise power spectral</i> <i>thermal noise power spectral density</i> white noise <i>white noise interference</i> <i>equivalent white noise</i> <i>equivalent white noise interference</i> <i>white noise interference source</i> noise amplifier(s) <i>low noise amplifier</i> <i>noise amplifier LNA</i> <i>low noise amplifier LNA</i> noise process <i>gated noise process</i> noise ratio(s) <i>thermal noise ratio</i> <i>noise ratio CNR</i> <i>noise ratio OSNR</i> <i>noise ratios SNRS</i> <i>signal to noise ratio(s)</i> <i>noise ration SNR degradation</i> <i>signal to noise ratio SNR</i></p>	<p><i>DB noise figure</i> <i>final noise figure</i> <i>low noise figure</i> <i>noise floor</i> <i>phase noise floor</i> <i>noise jamming</i> <i>radar noise jamming</i> <i>noise jamming techniques</i> <i>noise level(s)</i> <i>low noise levels</i> noise power <i>thermal noise power</i> <i>background noise power</i> <i>noise power spectral</i> <i>noise power spectral density</i> <i>thermal noise power spectral</i> <i>thermal noise power spectral density</i> noise source(s) <i>device noise sources</i> <i>dominant noise source</i></p>
--	--

4.3 Colocados contiguos de las combinaciones especializadas.

PRE +	COMBINACIÓN	+ POST
<i>including, low-frequency, represents, non-Gaussian</i>	<i>ambient noise</i>	<i>tends</i>
<i>reduces, high, ambient, end's, considering, active, signal-to-, differing, minimizes, radio</i>	<i>background noise</i>	<i>compared, level(s), power, ratio</i>
<i>wide-, in-, out-of-, broad</i>	<i>band noise</i>	<i>-spectral</i>
<i>additive, unknown</i>	<i>colored noise</i>	<i>case, input, signal</i>
	<i>crosstalk noise</i>	<i>avoidance</i>
	<i>external noise</i>	<i>environment, introduced, sources</i>
<i>low-frequency, name</i>	<i>flicker noise</i>	<i>generally, hints, intrudes</i>
<i>white, sub-, non-, additive, coloured, assuming, marginally, supplementary, fractional, zero-mean, complex</i>	<i>Gaussian noise</i>	<i>AWGN, channel(s), contribution, floor, impact, involve, level, ratio, samples, sequence, uncorrelated, vector</i>
<i>considerable, plus, frequency, require</i>	<i>impulsive noise</i>	<i>averaged, characterization, environments, due to, power, process, removal, sequence, show</i>
<i>low, excessive, chain, large, aggregate, integrated, VCO, typical, calculate, PLL</i>	<i>phase noise</i>	<i>according, appears, characteristics, degradation(s), floor, mask, performance, phase, random, relevant, spectra</i>
<i>long, optimized, PN</i>	<i>pseudo noise</i>	<i>PN, sequence, control</i>
<i>ADC, eliminate, reducing, sensor</i>	<i>quantization noise</i>	<i>tends</i>
<i>effectiveness</i>	<i>radar noise</i>	<i>jamming</i>
<i>disable, add, simple, pseudo, adding, level, added, containing, simulated</i>	<i>random noise</i>	<i>emissions, checkbox, level, -like, models, options, PN, reduces</i>
<i>current, noise, photodiode, effective, average</i>	<i>shot noise</i>	<i>current, due to, increases, level, power</i>
<i>background, derived, measured, radiometer</i>	<i>sky noise</i>	<i>temperature</i>
<i>ever-present, receiver, transmitter's, additive, output, including, representing</i>	<i>thermal noise</i>	<i>density, floor, leaking, limit, power, ratio, refers, relative, signal</i>
<i>produce, called, additive, equivalent, in-band, represents,</i>	<i>white noise</i>	<i>filtered, input, interference, rapidly, sequence, shown, signal,</i>

overwhelming, analog, using, Gaussian		source, speech-shaped, expansion, random
low	noise amplifier	designs, LNA , block, using, converters
AC, cyclostationary, interconnect, nonlinear, numerical, offer	noise analysis	available, method
adaptive, PM, phase-	noise cancellation	circuit, technique, strategy
Gaussian , greater, lower, transistor's, total, aliased, galactic	noise contribution	comes
collector, input, Leff, shot , large	noise current(s)	generator , density , power, associated, induced
causes, external, high, impulsive	noise enhancement	effect, refers
allowing, constant, devices', final, link, LAN's, low, lower, lowest, minimum, optimum, prototype, random , receive, receiver, smaller	noise figure(s)	lower, refers, circles, changes, measurement, test-set, calculator, meter, NF
spatial-temporal, more	noise filtering	
effective, Gaussian , phase, thermal , low	noise floor	slightly, combine
current, voltage	noise generator	tests
high, better, higher, substrate, adequate , improved	noise immunity	
common, radar, modulated	noise jamming	techniques, system, generation, usually, attempts
combined, folded , Gaussian , inherent , low, overall, random , shot , background , same, high, nonzero, different, lower, low	noise level(s)	rises
high, low, percentage, reduced, low-state, adequate , maintain, speed, better, narrow	noise margin(s)	NM, incidences, faults, problems, comes, performance
NMOS, spot	noise measurements	experiments, allows, techniques, reveals
class, Gaussian , LNA , phase, poor, low	noise performance	figures
background , combined, effective, fundamental, high, net, same, shot , thermal , carrier-to-, total, impulsive , receiver	noise power	due to, spectral , associated, ratio
gated, impulsive	noise process	occurs
background , Gaussian , interference , signal, thermal , carrier-to-, signal-to-, quantization	noise ratio(s)	SNR , level, OSNR , sets, CNR , criteria, requirement, situations, required, performance, PSNR , SQNR , calculator
possible, traditional, image	noise reduction	features, problem
perceptual	noise shaping	converters
low-level, white , narrowband , thermal , varying, Gaussian , various	noise signals	interaction, lends
dominant, fundamental, observed, same, white , device external, various, internal	noise source(s)	
wide-band, RTI, voltage	noise spectral	density
Sky , standard	noise temperature	attenuation, levels, results
accurate, instrument, unknown	noise variance	estimate(s), estimation
open-circuit, referred, mean-square(d), impulse , internal	noise voltage(s)	depended, $\sqrt{2}$, generator

3. Familia de *line*. Forma analizada: LINEAR

1. FRECUENCIA.

KEYWORD	Frec. Telec	Frec. Lacell	Ratio	Término, Chung	KEYNESS	P valor
90. LINEAR	1.590	276	21,8790543	NO	3.552,30	0
Miembros relacionados:	F Telec	F Gene	Ratio	Término	Keyness	P valor
NONLINEAR	402	7	218,106454	ESPECIFICO	1.193,30	0
ONLINE	937	790	4,50456651	NO	926,6	0
LINE / West	2.805	6.306	1,68934909	NO	495,5	0
LINES	1.240	2.173	2,16721732	NO	431,9	0
OFFLINE	131	12	41,4601199	NO	334	0
LINEARITY	126	17	28,1489453	NO	298,8	0
LINEARLY	104	15	26,3319235	NO	243	0
NONLINEARITIES	43	0	inf/esp	inf/esp	134,9	0
NONLINEARITY	41	1	155,712817	ESPECIFICO	119,6	0
LINEARIZATION	22	0	inf/esp	inf/esp	69	0
LINEWIDTH	20	0	inf/esp	inf/esp	62,7	0
BASELINE	66	82	3,05682507	NO	41,9	0
LINEARIZED	13	0	inf/esp	inf/esp	40,8	0
NEWLINE	15	2	28,4840518	NO	35,7	0
OUTLINE	161	394	1,55192296	NO	20,7	0,000005
TIMELINE	23	23	3,79787358	NO	19,1	0,000012
TIMELINESS	14	8	6,64627876	NO	18,8	0,000014
OLINER	5	0	inf/esp	inf/esp	15,7	0,000075
PHONELINE	5	0	inf/esp	inf/esp	15,7	0,000075
PHONELINES	5	0	inf/esp	inf/esp	15,7	0,000075
BASELINED	4	0	inf/esp	inf/esp	12,5	0,000397
Familia técnica: No, estadística y formalmente.						

2. DISTRIBUCIÓN.

Distribución por áreas	Distribución Keyness	Keykeyword	Nº de textos	Porcentaje
9	1,4,7	LINEAR	158	9,55%
<i>Dispersion plot</i>				
Area	Words	Hits	Per 1,000	
7 Systems	307.691	215	0.70	
4 Signal proc.	580.936	377	0.65	
1 Electronics	722.823	381	0.53	
2 Ar. Comp	329.643	139	0.42	
081 Esp. Sign	867.208	268	0.31	
5 Materials	101.241	26	0.26	
3 Telematics	1.205.064	130	0.11	
082 Esp. Tele	997.727	36	0.04	
6 Business	373.079	12	0.03	

3. COLOCACIÓN.

3.1 Colocados.

N	Collocates	TOTAL	LEFT	RIGHT	L5	L4	L3	L2	L1	*	R1	R2	R3	R4	R5
1	THE	915	508	407	112	106	79	84	127	0	0	61	140	100	106
2	OF	576	325	251	51	48	50	83	93	0	0	109	44	54	44
3	A	559	424	135	37	31	52	75	229	0	0	12	39	46	38
4	AND	462	190	272	41	32	32	52	33	0	65	66	41	52	48
5	IS	307	184	123	27	46	34	53	24	0	1	37	36	25	24
6	IN	292	158	134	38	37	23	35	25	0	25	30	27	25	27
7	TO	285	167	118	22	25	36	52	32	0	2	25	14	34	43
8	NON	264	236	28	1	3	1	1	230	0	2	15	3	7	1
9	FOR	197	112	85	18	20	14	22	38	0	2	20	17	24	22
10	WITH	139	58	81	6	9	12	17	14	0	11	27	19	11	13
11	ARE	138	70	68	11	17	15	17	10	0	0	16	16	14	22
12	AS	137	81	56	14	14	10	26	17	0	5	7	20	15	9
13	SYSTEMS	131	45	86	12	3	10	9	11	0	27	18	20	14	7
14	BE	114	73	41	18	23	13	6	13	0	0	2	15	13	11
15	TIME	101	34	67	8	5	7	8	6	0	41	12	5	3	6
16	BY	95	63	32	7	17	11	18	10	0	1	3	15	8	5
17	THAT	89	51	38	13	15	10	11	2	0	0	7	9	7	15
18	ON	82	52	30	7	7	5	18	15	0	2	8	3	11	6
19	SYSTEM	76	26	50	5	4	6	11	0	0	26	8	6	5	5
20	THIS	76	47	29	16	15	7	3	6	0	0	10	7	8	4
21	AN	64	35	29	10	6	5	14	0	0	0	4	9	7	9
22	ANALYSIS	64	19	45	1	5	3	6	4	0	13	14	8	7	3
23	PHASE	64	9	55	2	1	4	2	0	0	48	3	1	1	2
24	WHICH	62	23	39	7	4	5	7	0	0	0	11	9	9	10
25	CAN	61	28	33	11	12	4	1	0	0	0	12	6	7	8
26	MODEL	59	12	47	2	3	4	3	0	0	27	8	4	2	6
27	WE	59	27	32	9	10	5	3	0	0	1	8	12	2	9
28	IT	56	21	35	9	6	3	2	1	0	0	14	7	8	6
29	NONLINEAR	55	12	43	5	4	1	2	0	0	0	26	9	5	3
30	OR	54	27	27	3	6	4	8	6	0	8	7	5	3	4
31	CONTROL	50	13	37	2	3	6	2	0	0	12	7	10	4	4
32	INTERPOLATION	48	13	35	5	1	1	6	0	0	30	3	2	0	0
33	BASED	46	29	17	7	3	8	9	2	0	0	9	4	2	2
34	FILTERS	46	13	33	3	1	5	1	3	0	10	5	11	1	6
35	INVARIANT	46	9	37	1	2	0	3	3	0	0	26	8	2	1

3.2 Colocados significativos.

A) Colocados significativos en el corpus técnico.

Nº	Colocados	Z	T	MI	Nº	Colocados	Z	T	MI
1	NONLINEAR	50,102	7,2604	5,5734	52	APPROXIMATION	7,9579	2,6638	3,1577
2	INTERPOLATION	80,481	6,8776	7,0973	53	FORMULATION	10,677	2,6537	4,0168
3	INVARIANT	60,707	6,6997	6,3593	54	REVENUE	7,5558	2,6351	3,0394
4	FILTERS	22,841	6,271	3,7294	55	MBPC	32,433	2,6284	7,2504
5	PREDICTION	39,993	5,8707	5,5362	56	MULTIVARIABLE	23,961	2,6143	6,3924
6	EQUATIONS	26,402	5,6456	4,4508	57	REPRESENTATIONS	9,0748	2,5968	3,6102
7	ALGEBRA	47,029	5,1342	6,3906	58	ICS	9,0282	2,5948	3,5976
8	DISCRETE	16,518	4,9794	3,4597	59	POLARIZATION	8,675	2,5785	3,5006
9	EXTRAPOLATION	67,147	4,6677	7,6931	60	DISPLACEMENT	11,949	2,5274	4,4822
10	REGRESSION	35,958	4,5105	5,9899	61	PERFECTLY	11,44	2,5176	4,3678

11	MULTIUSER	20,428	4,4662	4,3868	62	IMPULSE	7,2561	2,4942	3,0811
12	GEOMETRICALLY	61,013	4,4484	7,5555	63	AMPLIFIERS	7,0659	2,48	3,021
13	PREDICTIVE	31,86	4,3873	5,7206	64	SPACES	9,5779	2,4698	3,9106
14	PIECEWISE	61,015	4,2223	7,7061	65	TRANSFORMS	9,478	2,4666	3,8841
15	GRADIENTS	45,736	4,2067	6,8851	66	MATRICES	9,0731	2,4524	3,7747
16	COMBINATION	12,896	4,1942	3,2407	67	DETERMINISTIC	8,7894	2,4416	3,6958
17	FIR	20,015	4,1697	4,5261	68	LPC	32,127	2,4354	7,4431
18	MATRIX	12,384	4,0841	3,2006	69	ENVELOPE	8,5775	2,4329	3,6357
19	INTEGER	15,693	4,0662	3,8966	70	CELP	27,769	2,4307	7,028
20	TEMPORAL	14,258	4,0135	3,6576	71	CURVES	7,6576	2,3884	3,3617
21	MMSE	41,414	3,9634	6,7706	72	LTI	13,343	2,3721	4,9836
22	TRANSFORM	11,514	3,8672	3,148	73	LOGARITHMIC	11,647	2,3498	4,6186
23	OPTIMIZED	13,338	3,6933	3,705	74	NOTATION	6,9387	2,3438	3,1315
24	EXCITED	25,716	3,6656	5,6211	75	CART	10,32	2,3251	4,3001
25	ADAPTIVE	10,855	3,6555	3,1404	76	GRADIENT	9,8631	2,3146	4,1825
26	CONVOLUTION	22,479	3,6434	5,2504	77	TRANSCONDUCTANCE	9,8156	2,3134	4,1701
27	CIRCULAR	20,311	3,6226	4,9743	78	FACILITATES	9,7686	2,3122	4,1577
28	DETECTORS	19,394	3,6119	4,8496	79	ADDITIVE	9,7686	2,3122	4,1577
29	FILTERING	11,163	3,587	3,2757	80	OBTAINING	9,3703	2,3017	4,0508
30	LIGHTWAVE	43,985	3,5816	7,2366	81	CANCELLATION	8,2925	2,2665	3,7426
31	BEHAVIOUR	12,283	3,5495	3,5818	82	SMOOTH	8,0542	2,2571	3,6705
32	QUADRATIC	26,465	3,541	5,8037	83	EXPONENTIAL	7,3355	2,2243	3,4431
33	OPTIMUM	13,68	3,3848	4,0298	84	GPC	22,402	2,2142	6,6775
34	SHELLS	25,223	3,2612	5,9025	85	OPTICS	7,0088	2,2067	3,3346
35	EQUALIZER	20,091	3,2309	5,2732	86	DYNAMICS	6,9883	2,2055	3,3276
36	DIFFERENTIAL	9,6841	3,2095	3,1864	87	DECORRELATOR	17,673	2,2014	6,0101
37	IIR	16,276	3,1893	4,7028	88	DISCONTINUOUS	14,602	2,186	5,4796
38	COEFFICIENT	10,407	3,1473	3,4506	89	WIENER	12,064	2,1641	4,9576
39	MAGNITUDE	8,6703	3,0386	3,0252	90	TREATMENT	6,3123	2,1621	3,0914
40	BUCKLING	52,925	2,9904	8,2911	91	EXTERNALLY	11,046	2,1513	4,7206
41	CHARACTERISTIC	9,5285	2,99	3,3441	92	APPROXIMATIONS	9,6233	2,1268	4,3556
42	POLE	11,432	2,9515	3,907	93	WISE	9,5614	2,1256	4,3387
43	CONTROLLERS	10,063	2,8997	3,5901	94	QUASI	8,6001	2,1024	4,0646
44	ALGEBRAIC	15,372	2,8937	4,8186	95	DIFFRACTION	8,4631	2,0986	4,0235
45	GENERALIZED	11,907	2,8305	4,1454	96	INTEGRATOR	7,8903	2,0806	3,8461
46	APPROXIMATE	10,665	2,7941	3,8648	97	SPATIALLY	7,853	2,0793	3,8343
47	MIXED	8,1283	2,7897	3,0855	98	LINEARITY	7,7083	2,0742	3,7877
48	SQUARE	8,0414	2,7834	3,0611	99	AMPLIFICATION	7,2493	2,0562	3,6357
49	ELASTIC	20,516	2,7766	5,7706	100	BLIND	6,8781	2,0395	3,5076
50	EQUALIZERS	19,916	2,7736	5,6882	101	EXPENSE	6,5732	2,024	3,3987
51	SPEEDUP	14,162	2,7238	4,7566	102				

B) Colocados significativos en el corpus general.

N°	Colocados	Z	T	MI
1	VISCOELASTICITY	240,18	3,46	12,23
2	MEASUREMENT	98,95	4,12	9,17
3	DIFFERENTIAL	33,92	2,23	7,86
4	CIRC	77,49	2,64	9,75

3.3 Definición del nodo en el diccionario.

1. Significado de *linear* en la *General Service List* (West, 1953):

Line (noun):

1. (*string, wire*): *Telegraph line, fishing line.*
2. (*mark*) *A pencil line. A thing line of paint.*
3. (*showing direction, tendency, type*) *Line of force, sight, march. Line of thought, business.*
4. (*showing arrangement, edge, or contour*) *Stand inline, in one line. Bring ino line.*
5. (*showing a limit*) *Must draw a line somewhere. Frontier line.*
6. (*of ships and railway*) *The Cunard Line. On the railway line.*
7. (*military uses*) *In the front line; line of battle.*
8. (*print, poetry, handwriting*) *A line of print, of poetry.*

2. Definición de *Linear* en el *Advanced Learner's Dictionary* de Cambridge (2003):

Linear (adj):

1. (*lines*) *Consisting of or to do with lines: a linear diagram.*
2. (*length*) *Relating to length, rather than area or volume: linear measurement.*
3. (*connection*)
 1. *Involving a series of events or thoughts in which one follows another one directly: These mental exercises are designed to break linear thinking habits and encourage the creativity that is needed for innovation.*
 2. *Describes a relationship between two things that is direct or clear: Is there a linear relationship between salaries and productivity?*

3. Definición de *linear* en el diccionario especializado Webster:

Linear (adj):

1. (*mathematics*) *directly proportional.*
2. *Of or in or along or relating to a line; involving a single dimension; "a linear foot".*
3. (*electronics*) *of a circuit or device having an output that is proportional to the input; "analogue device"; "linear amplifier".*
4. *Of a leaf shape; long and narrow.*
5. *Measured lengthwise; "cost of lumber per running foot".*

Specialty Definition:

In aerospace:

1. *Of or pertaining to a line.*
2. *Having a relation such that a change in one quantity is accompanied by an exactly proportional change in a related quantity, such as input and output of electronic equipment.*

In computing:

Qualifies a circuit having an output that varies in direct proportion to the input.

In post & telecom:

Pertaining to an operating characteristic, a device, an electrical network or a transmission medium satisfying the superposition principle according to which, within a specified working range at a specified output port, the signal resulting from an arbitrary input signal, applied to a given port is such that, when the input signal is increased by a numerical factor, the output signal is also increased by the same factor.

3.4 Colocados contiguos.

a) Colocado + <i>linear</i> .		
Total colocados seleccionados: 21	Frecuencia máxima: 230	Frecuencia mínima: 4
<ul style="list-style-type: none"> - Sustantivos, total 3: <i>integrer 17, output 6, order 4.</i> - Adjetivos, total 8: <i>local 24, approximate 8, general 6, optimum 6, standard 5, variable 5, wise 5, optimal 4.</i> - Participios verbales que actúan como adjetivos, total 1: <i>excited 12.</i> - Participios verbales, total 3: <i>using 10, generalizad 6, called 4.</i> 		

<ul style="list-style-type: none"> - Adverbios, total 5: <i>piecewise</i> 14, <i>approximately</i> 8, <i>perfectly</i> 5, <i>essentially</i> 4, <i>externally</i> 4. - Prefijos, total 2: <i>non</i> 230, <i>quasi</i>- 4. - Formas abreviadas, total 1: <i>mmse</i> 4. 		
b) <i>Linear</i> + colocado.		
Total colocados seleccionados: 41	Frecuencia máxima: 48	Frecuencia mínima: 6
<ul style="list-style-type: none"> - Sustantivos, total 33: <i>phase</i> 48, <i>time</i> 41, <i>prediction</i> 32, <i>interpolation</i> 30, <i>systems</i> 27, <i>algebra</i> 25, <i>combination</i> 17, <i>function</i> 16, <i>circuit</i> 16/circuits 11, <i>models</i> 15, <i>regression</i> 13, <i>lightwave</i> 13, <i>control</i> 12, <i>array</i> 12, <i>sequence</i> 11, <i>range</i> 11, <i>relationship</i> 10, <i>parameter</i> 10, <i>equalizer</i> 10, <i>current</i> 10, <i>behaviour</i> 10, <i>region</i> 9, <i>multiuser</i> 9, <i>filter</i> 9, <i>transformation</i> 8, <i>program</i> 8, <i>network</i> 8, <i>effects</i> 8, <i>detection</i> 8, <i>transfer</i> 7, <i>process</i> 6, <i>speedup</i> 6. - Adjetivos, total 6: <i>temporal</i> 11, <i>constant</i> 9, <i>predictive</i> 8, <i>analog</i> 8, <i>quadratic</i> 7, <i>elastic</i> 7. - Participios verbales, total 2: <i>programming</i> 19, <i>buckling</i> 9. 		

3.5 Identificación de colocados contiguos especializados en las líneas de concordancia.

ar FE analysis [9] is carried out to investigate the non-linear buckling loads of the optimized shells. The met e non-optimized shells. It should be noted that the non-linear buckling load due to the snap-through is mainly istributions of the shell have been determined, the non-linear buckling loads of the initial and optimized shells lower than that of the standard shell. However, the non-linear buckling load of the optimum shell obtained usi ptimized shell which has the value of 0.581. The non-linear buckling load parameter of the ‘thickness-optimi From the numerical results shown in Fig. 12, the non-linear buckling load parameter of the ‘shape-optimized sed on arc length control. The tests show that the non-linear buckling load of shells optimized using thicknes the standard non-optimized shells. In addition, the non-linear buckling load of a shell optimized using mixed s lysis. 3.1. Introduction. As mentioned earlier, the non-linear buckling load of the optimized shell obtained fro

Concordancias de linear buckling.

a) Colocado + linear:

approximate linear	non linear	quasi linear
externally linear	order linear	standard linear
general linear	output linear	variable linear
integer linear	perfectly linear	wise linear
local linear	piecewise linear	

b) *Linear* + colocado:

linear algebra	linear equalizer	linear process
linear algebraic	linear equations	linear program
linear analog	linear filter	linear programming
linear approximation	linear function(s)	linear quadratic
linear array	linear increase	linear range
linear behaviour	linear interpolation	linear region
linear buckling	linear lightwave	linear regression
linear circuit(s)	linear mechanical	linear relationship
linear combination	linear model(s)	linear sequence
linear constant	linear multiuser	linear speedup
linear control	linear multivariable	linear system(s)
linear convolution	linear network	linear temporal
linear current	linear parameter	linear time
linear detection	linear phase	linear transfer
linear effects	linear prediction	linear transformation
linear elastic	linear predictive	

4. AGRUPACIONES LÉXICAS O CLUSTERS.

4.1 Distribución de clusters.

Combinaciones	Volumen	Frecuencia máxima
2-cluster	206	205
3-cluster	175	6
4-cluster	123	10
5-cluster	71	6
6-cluster	46	6

4.2 Inclusión de combinaciones especializadas en unidades más amplias.

<p><i>excited linear</i> code excited linear prediction general linear general linear circuit integer linear integer linear program integer linear problem an integer linear programming local linear local linear models local linear MBPC local linear MBPC controllers mmse linear mmse linear multiuser mmse linear multiuser detection non linear non linear effects linear phase linear phase filters linear phase property non-linear phase linear phase analysis linear phase response generalized linear phase property linear systems with generalized linear phase linear phase analysis and synthesis filter linear lightwave linear lightwave network linear control non-linear control systems linear sequence linear sequence number linear detection linear detection techniques non-linear detection linear function(s) non-linear function(s)</p>	<p>linear time linear time invariant linear prediction excitation linear prediction code excited linear prediction mixed excitation linear prediction linear system(s) non-linear system(s) linear system theory time linear systems linear model(s) linear modelling local linear models linear modeling errors linear programming linear programming formulation an integer linear programming linear equation(s) non-linear equation(s) system of linear equations a set of linear equations linear circuit(s) non-linear circuit(s) general linear circuit linear regression linear regression models linear analog linear analog circuits linear buckling non-linear buckling load linear effects non-linear effects linear filter(s) linear filtering optimum linear filters linear parameter linear parameter gradients linear transfer linear transfer function</p>
--	---

4.3 Colocados contiguos de las combinaciones especializadas.

PRE +	COMBINACIÓN	+ POST
code, sum	<i>excited linear</i>	prediction
	<i>general linear</i>	circuit, model
<i>mixed, using</i>	<i>integer linear</i>	MILP, programming, program
several, three, LOLIMOT, following,	<i>local linear</i>	controllers, dynamic, GPC, MBPC, model(s), model-based
<i>adaptive,</i> minimum-mean-square-error	<i>mmse linear</i>	equaliser, multiuser
highly, subject to, geometrically, being, investigate, produce, strongly, using, methods of,	<i>non linear</i>	signal, effects, errors, function(s), Kerr, means, medium, system, transformation, adaptive, algorithms, amplifier, analysis,

initially, larger, keep, minimize, sensitive to, discontinuous , single, simultaneous, linearized, three, pre- spatially , residual, present, ordinary, pole-cart , counter-intuitive, fundamental, IIR , real, generalised, MIMO, special, strongly		applications, arrival, autoregressive, behaviour , buckling , characteristic , circuit(s), classifiers, compensation, components, control, current, DC, degradations, detection, development, device (s), discontinuous , discrete , distortion, documents, dynamic, effects, element(s), equation(s), fashion, FE, filtered, gradients , model(s), nature, navigation, object, operations, optic(s) , optical, parametric, partial, phase, recursion(s), quadratic , region, satellite, self-focusing, shell, solution, standpoint, squares , transformation(s), transponder, TWTA, way, weighted, interpolation
fault-tolerant	linear analog	circuits, network, operation,
non-	linear buckling	load(s)
advanced, certain, driven, general, ideal, non-, systems, noise, externally , high-end	linear circuit(s)	element(s), simulation, simulator, driven, analysis
conventional, non-, robust, stiff, traditional	linear control	design, systems, theory
non-	linear detection	technique(s), approaches, methods, isolation
non-	linear effects	give, prevent
non-, a set of, fairly, precisely	linear equation(s)	augmented, systems, using
time-invariant, optimum, FIR , non-, adaptive , using, possible	linear filter	representing, given, frames, methods
exactly, non-, piecewise , continuous	linear function(s)	followed
called, preventing,	linear lightwave	network(s)
non-, use(s), using, performing, piecewise-	linear interpolation	rule, extrapolation
general, local, order, non-, smooth	linear model(s)	based, trees, predictive , errors, contained, present
predict,	linear parameter	varying, gradients
almost, approximate (ly), design, flat, generalized , IFIR, no, perfect, prevents, called, perfectly , pure, non-	linear phase	response, FIR , filter (s), polyphase, IIR , M-band, property, characteristic , recursive, means, restriction, requirement, analysis
backward, contrasting, DFT, excitation, excited , approximate , rudimentary	linear prediction	MELP, ACELP, CELP , WAP, codecs, analysis, theory, -error
compact, derived, discusses, integer , non-, sequential, compact	linear programming	based, approach, model, formulation , ILP, based, NLP, SLP
calculated, electromagnetic, exploiting, multiple	linear regression	approximation , analysis, model(s), lines, EMLRM
non-, multiple-output, primarily, sparse, baseband, pertinent, uncertain, discrete-time , continuous	linear system(s)	theory, seems, provides, models, solvers, identification, analysis, NPID, concepts, vector
new, have, finite-dimensional, functions	linear time	parameterisation, version, reverse, invariant , complexity, harmonic, -variant
non-	linear transfer	function(s), matrix , characteristic ,

4. Familia de *programme*. Forma analizada: PROGRAMMING

1. FRECUENCIA.

KEYWORD	Frec. Telec	Frec. Lacell	Ratio	Término, Chung	KEYNESS	P valor
102. PROGRAMMING	1.560	326	18,1738736	NO	3.308,90	0
Miembros relacionados:	F Telec	F Gene	Ratio	Término	Keyness	P valor
PROGRAM / West (programme)	2.945	2.418	4,62561525	NO	2.984,90	0
PROGRAMS	1.293	1.261	3,89425102	NO	1.104,80	0
PROGRAMMABLE	364	16	86,4016239	ESPECIFICO	1.016,50	0
PROGRAMMERS	254	65	14,8409214	NO	504,5	0
PROGRAMMER	267	88	11,5230937	NO	481	0
SUBPROGRAM	60	0	inf/esp	inf/esp	188,2	0
PROGRAMMED	118	105	4,2680865	NO	110,8	0
PROGRAMMABILITY	25	0	inf/esp	inf/esp	78,4	0
SUBPROGRAMS	15	0	inf/esp	inf/esp	47	0
MULTIPROGRAMMED	10	0	inf/esp	inf/esp	31,4	0
REPROGRAMMING	12	2	22,7872415	NO	27,1	0
REPROGRAMMED	10	1	37,9787358	NO	25,1	0,000001
REPROGRAMMABLE	7	0	inf/esp	inf/esp	22	0,000003
PROGRAMMATICALLY	6	0	inf/esp	inf/esp	18,8	0,000014
MULTIPROGRAMME	5	0	inf/esp	inf/esp	15,7	0,000075

Familia técnica: No, estadística y formalmente.

2. DISTRIBUCIÓN.

Distribución por áreas	Distribución Keyness	Keykeyword	Nº de textos	Porcentaje
9	0	PROGRAM	97	5,86%
<i>Dispersion plot</i>				
Area	Words	Hits	Per 1,000	
7 Systems	307.691	287	0.93	
5 Materials	101.241	92	0.91	
3 Telematics	1.205.064	591	0.49	
2 Ar. Comp	329.643	146	0.44	
4 Signal proc.	580.936	161	0.28	
1 Electronics	722.823	90	0.12	
081 Esp. Sign	867.208	112	0.11	
082 Esp. Tele	997.727	59	0.07	
6 Business	373.079	19	0.05	

3. COLOCACIÓN.

3.1 Colocados.

N	Collocates	TOTAL	LEFT	RIGHT	L5	L4	L3	L2	L1	*	R1	R2	R3	R4	R5
1	THE	748	396	352	74	94	56	117	55	0	14	107	88	72	71
2	OF	545	410	135	57	47	74	143	89	0	8	13	29	36	49
3	AND	493	236	257	38	39	50	64	45	0	46	80	36	49	46
4	A	403	250	153	37	58	35	62	58	0	6	40	35	41	31
5	IN	373	185	188	36	27	49	44	29	0	58	52	27	21	30
6	TO	331	204	127	39	44	56	30	35	0	4	33	27	21	42

7	IS	278	84	194	34	20	21	9	0	0	46	34	50	40	24
8	LANGUAGE	243	24	219	9	5	3	4	3	0	197	2	8	4	8
9	FOR	238	116	122	20	9	28	26	33	0	10	41	23	22	26
10	LANGUAGES	159	12	147	0	5	5	1	1	0	130	6	5	2	4
11	ARE	126	53	73	20	10	13	7	3	0	7	13	20	15	18
12	THAT	123	40	83	15	10	10	4	1	0	10	34	13	16	10
13	WITH	114	52	62	13	11	11	11	6	0	24	11	9	9	9
14	AS	112	49	63	14	11	14	7	3	0	8	10	25	12	8
15	ON	112	57	55	11	11	12	21	2	0	11	11	14	6	13
16	FUNCTIONAL	98	97	1	2	4	4	18	69	0	0	0	0	1	0
17	AN	95	48	47	14	9	9	16	0	0	1	12	15	11	8
18	BE	86	39	47	13	14	8	3	1	0	1	7	11	14	14
19	MODEL	80	15	65	1	3	4	7	0	0	51	3	3	4	4
20	IT	79	17	62	6	7	2	0	2	0	1	21	17	16	7
21	THIS	78	29	49	7	7	9	6	0	0	0	19	10	11	9
22	JAVA	73	45	28	2	6	3	1	33	0	0	16	0	7	5
23	C	72	37	35	4	3	1	3	26	0	0	8	12	9	6
24	CONCURRENT	69	62	7	4	4	4	3	47	0	1	0	4	2	0
25	BY	67	35	32	5	8	10	6	6	0	4	5	11	1	11
26	OR	65	26	39	5	7	9	5	0	0	10	7	6	8	8
27	PARALLEL	63	52	11	1	0	5	3	43	0	1	3	2	3	2
28	CAN	61	20	41	9	6	1	3	1	0	8	5	9	12	7
29	FROM	55	34	21	7	7	7	9	4	0	1	3	3	3	11
30	WE	55	21	34	9	9	2	1	0	0	0	13	8	5	8
31	USING	54	33	21	8	4	15	6	0	0	8	3	1	3	6
32	LOGIC	53	46	7	0	2	1	1	42	0	2	3	0	1	1
33	OBJECT	50	38	12	5	0	5	28	0	0	2	1	1	6	2
34	NOT	49	27	22	10	6	8	2	1	0	0	4	7	6	5
35	APPLICATION	48	36	12	4	3	3	3	23	0	0	1	6	4	1

3.2 Colocados significativos.

A) Colocados significativos en el corpus técnico.

Nº	Colocados	Z	T	MI	Nº	Colocados	Z	T	MI
1	GENERATIVE	45,60	3,72	7,23	37	LABVIEW	22,54	5,26	4,20
2	ILP	25,03	2,43	6,73	38	PARAMETRIC	9,60	2,31	4,11
3	PARADIGMS	36,56	3,57	6,71	39	ADA	12,29	2,98	4,09
4	LANGUAGES	120,15	12,47	6,54	40	SEQUENTIAL	18,23	4,42	4,09
5	INTRODUCTORY	32,51	4,06	6,00	41	BASICS	9,23	2,30	4,01
6	PASCAL	28,46	3,68	5,90	42	TEACHING	9,72	2,47	3,95
7	JOURNAL	38,79	5,11	5,85	43	CONVENTIONAL	21,06	5,44	3,91
8	PREFACE	16,62	2,20	5,84	44	EASE	11,25	2,95	3,87
9	GP	18,87	2,60	5,72	45	POOL	10,02	2,63	3,86
10	OVERLOADING	15,60	2,19	5,66	46	PREPARATION	8,64	2,28	3,85
11	REWRITING	21,29	3,10	5,56	47	CNC	8,58	2,28	3,83
12	CONCURRENT	54,48	8,12	5,49	48	CORBA	11,54	3,08	3,81
13	FUNCTIONAL	62,78	9,66	5,40	49	GENERIC	15,36	4,15	3,78
14	PROCEDURAL	16,61	2,58	5,37	50	SYNTAX	12,36	3,34	3,77
15	INDUCTIVE	17,70	2,76	5,36	51	MATHEMATICAL	11,67	3,31	3,63
16	TEACH	15,29	2,39	5,36	52	FORMULATION	7,92	2,25	3,63
17	PARADIGM	29,90	4,77	5,29	53	VISUAL	10,05	2,90	3,59
18	EARTH'S	13,08	2,17	5,18	54	INTEGER	12,29	3,55	3,58
19	IMPERATIVE	12,30	2,17	5,01	55	CONCURRENCY	8,82	2,59	3,54
20	ABSTRACTIONS	13,48	2,37	5,01	56	STRUCTURED	9,32	2,74	3,53

21	EXERCISES	19,77	3,49	5,00	57	SOLVED	7,47	2,23	3,49
22	QUADRATIC	14,17	2,56	4,94	58	DRIVEN	12,12	3,73	3,40
23	GRAPHICAL	23,92	4,43	4,87	59	ASSEMBLY	9,24	2,86	3,38
24	MACHINING	11,07	2,15	4,73	60	EARTH	9,01	2,85	3,32
25	THREADED	15,53	3,04	4,71	61	PRINCIPLES	9,37	2,98	3,30
26	STYLE	24,52	4,81	4,70	62	THREADS	9,62	3,10	3,26
27	API	16,97	3,33	4,70	63	MATLAB	7,18	2,36	3,21
28	ORIENTED	31,18	6,36	4,59	64	STRATEGIES	11,12	3,67	3,20
29	FORTRAN	11,34	2,34	4,55	65	LEARN	10,39	3,45	3,18
30	CONSTRUCTS	18,50	3,94	4,47	66	PASSING	8,01	2,67	3,17
31	TEXTBOOK	9,99	2,13	4,45	67	CONSTRAINT	9,61	3,20	3,17
32	MECHANICS	9,99	2,13	4,45	68	THEOREM	7,02	2,35	3,16
33	EXTREME	12,06	2,69	4,33	69	MOTION	7,23	2,49	3,07
34	JANUARY	15,30	3,42	4,32	70	CONCEPTS	11,14	3,84	3,07
35	JAVA	35,56	8,10	4,27	71	ASPECT	6,57	2,32	3,01
36	GENETIC	9,20	2,12	4,24					

B) Colocados significativos en el corpus general.

Nº	Colocados	Z	T	MI
1	PARAMETRIC	208,12	3,32	11,94
2	MULTI	45,35	2,23	8,69
3	ORIENTED	46,23	2,44	8,48
4	INTERFACE	44,57	2,82	7,97
5	LOGIC	49,11	3,59	7,55
6	LINEAR	35,79	2,63	7,53
7	USER	28,50	2,22	7,36
8	PARALLEL	16,67	2,20	5,85
9	ADVANTAGES	15,59	2,19	5,66

3.3 Definición del nodo en el diccionario.

1. Significado de *programme* en la *General Service List* (West, 1953):

Programme (noun):

1. (printed paper) *Theatre programme.*
2. (scheme) *A very ambitious programme for our visit to London.*

Programme-/ *Programme-seller, et.*

2. Definición de *program(me)* y *programming* en el *Advanced Learner's Dictionary de Cambridge* (2003):

Program (noun):

A series of instructions which can be put into a computer in order to make it perform an operation: a computer program.

Program (verb):

To write a series of instructions which make a computer perform a particular operation.

Computer program is spelt **program in both British English and American English. In British English, **programme** is the spelling used for all other meanings. In American English, **program** is the spelling used for all meanings.*

Programme (noun):

1. (Broadcast) *a broadcast on television or radio: It's one of those arts programmes late at night.*
2. (Thin book) *a thin book or piece of paper giving information about a play or musical or sports event, usually bought at the theatre or place where the event happens: I looked in the programme to find out the actor's name.*
3. (Plan) *a plan of activities to be done or things to be achieved: The school offers an exciting and varied programme of social events.*

Programming (noun)

When someone writes computer programs.

3. Definición de *programming* en el diccionario especializado Webster:

Programming (noun):

1. *Setting an order and time for planned events.*
2. *Creating a sequence of instructions to enable the computer to do something.*

Specialty Definition:

In computing:

The art of debugging a blank sheet of paper (or, in these days of on-line editing, the art of debugging an empty file).

In electrical engineering:

The control of a power supply parameter, such as output voltage, by means of a control element of a signal.

3.4 Colocados contiguos.

a) Colocado + <i>programming</i> .		
Total colocados seleccionados: 37	Frecuencia máxima: 69	Frecuencia mínima: 6
<ul style="list-style-type: none"> - Sustantivos, total 13: <i>Java 33, application 23, level 21, Labview 15, systems 9, device 8, abstract 7, computer 7, earth 6, extreme 6, scheme 6, software 6, user 6.</i> - Adjetivos, total 20: <i>functional 69, concurrent 47, logic 42, conventional 22, dynamic 20, linear 20, generic 16, graphical 16, sequential 16, introductory 13, generative 10, mathematical 10, common 9, strategic 8, basic 7, powerful 7, different 6, good 6, new 6, quadratic 6.</i> - Participios verbales, total 3: <i>oriented 32, distributed 11, driven 10.</i> - Formas abreviadas, total 1: <i>C 26.</i> 		
b) <i>Programming</i> + colocado.		
Total colocados seleccionados: 15	Frecuencia máxima: 197	Frecuencia mínima: 7
<ul style="list-style-type: none"> - Sustantivos, total 13: <i>language 197/languages 130, model 51/models 18, environment 25/environments 15, techniques 21, interface 19/interfaces 12, paradigm 11/paradigms 9, style 10, constructs 9, tools 7.</i> - Participios verbales, total 1: <i>based 7.</i> 		

3.5 Identificación de colocados contiguos especializados en las líneas de concordancia.

ter programs. LEGO materials and the LabVIEW programming environment are the tools that allow logic programming paradigm. The EPSILON logic programming environment allows the structuring of s when our new JVM is executed on top of a multi-programming environment such as Solaris. The O peed; but for a theorem-prover developer, quality of programming environment is equally important. De e passing) and on the other hand on the Parallaxis programming environment in which program execut he Engineering of Concurrent Programs. A parallel programming environment must support the followi inistic execution is avoided. Extension of the pool programming environment to support global data is ing and distributed threads of execution. The target programming environment of GraphIcsla is the Pro an interpreter to execute the Java bytecode. Your programming environment might automate some of ew features of the next version of Microsoft's .NET programming environment, Visual Studio 2005. Wh f engineers and scientists, and as part of an open programming environment, provides an ideal platfo S, they have little difficulty moving into the diagram programming environment. Interestingly, however, t ach station is also equipped with the Vee graphical programming environment. The laboratory will acc ach station is also equipped with the Vee graphical programming environment. The laboratory will acc y the user, or system libraries that are a part of the programming environment. For example, procedure urse based on this textbook, I have used a different programming environment: CodeWarrior for Macinto done considerable work related to middleware and programming environments for mobile and context- ating data such as strings or image data. Common programming environments provide rules for referen

the most useful benefit of dataflow-based graphical **programming environments** for DSP is that carefull Submitted: 2002-07-24. Nowadays the integrated **programming environments** - Borland Delphi and B ndix, I'll give some information and opinion on Java **programming environments** for Windows, Macintosh e part of the program that contains the error. Most **programming environments** come with a debugger, significantly eased the design and construction of **programming environments** (PEs). Concurrently, ad RECTIONS. LEGO Engineer and RoboLab are two **programming environments** that extend the capabi built-in equational theories. To provide reasonable **programming environments**, compilation techniques struction of new languages and application-specific **programming environments**, and so on. As you b ou are using. See Appendix 2 for information about **programming environments**. In general, you just ha ptable time and space overhead. A foundation for **programming environments**. Recent work on gener fptr is NULL. Appendix 2: Some Notes on Java **Programming Environments**. EACH TIME I HAVE ograms end up being, in effect, application-specific **programming environments**. In fact, this style of pr dure calls and is considered to represent systems **programming environments**. The results are usuall

Concordancias de programming environment(s).

a) Colocado + programming:

<i>abstract programming</i>	extreme programming	mathematical programming
<i>application programming</i>	functional programming	oriented programming
<i>basic programming</i>	generative programming	<i>parallel programming</i>
<i>C programming</i>	generic programming	quadratic programming
<i>computer programming</i>	graphical programming	<i>scheme programming</i>
concurrent programming	introductory programming	sequential programming
conventional programming	Java programming	<i>software programming</i>
<i>device programming</i>	LabVIEW programming	<i>strategic programming</i>
<i>distributed programming</i>	<i>level programming</i>	<i>systems programming</i>
<i>dynamic programming</i>	<i>linear programming</i>	<i>user programming</i>
earth programming	<i>logic programming</i>	event-driven programming

b) Programming + colocado:

<i>programming based</i>	<i>programming paradigm(s)</i>
<i>programming environment(s)</i>	<i>programming style</i>
<i>programming interface(s)</i>	<i>programming techniques</i>
<i>programming language(s)</i>	<i>programming tools</i>
<i>programming model(s)</i>	

4. AGRUPACIONES LÉXICAS O CLUSTERS.

4.1 Distribución de clusters.

Combinaciones	Volumen	Frecuencia máxima
2-cluster	176	203
3-cluster	155	9
4-cluster	64	8
5-cluster	27	8
6-cluster	16	5

4.2 Inclusión de combinaciones especializadas en unidades más amplias.

<p>application programming <i>application programming interfaces</i> <i>application programming interface API</i> concurrent programming <i>principles of concurrent programming</i> conventional programming <i>conventional programming language</i> distributed programming <i>parallel and distributed programming</i> dynamic programming <i>dynamic programming algorithm</i> <i>dynamic programming equations</i></p>	<p>earth programming <i>earth programming model</i> functional programming <i>functional programming language</i> <i>functional programming paradigm</i> <i>the journal of functional programming</i> graphical programming <i>graphical programming language</i> Java programming <i>Java programming environments</i> level programming <i>level programming language(s)</i></p>
--	--

<p>logic programming inductive logic programming logic programming languages constraint logic programming control task with inductive logic programming</p> <p>oriented programming object oriented programming programming language(s) level programming language(s) graphical programming language logic programming language functional programming language parallel programming language conventional programming language powerful programming language programming language constructs sequential programming language ADA programming language common programming languages programming language developed programming language terminology purpose programming language C programming language and matlab knowledge of C programming language</p>	<p>linear programming nonlinear programming integer linear programming linear programming formulation quadratic programming sequential quadratic programming</p> <p>parallel programming parallel programming language parallel programming model(s) data parallel programming</p> <p>sequential programming sequential programming language conventional sequential programming</p> <p>programming environment(s) interactive programming environment Java programming environments</p> <p>programming interface(s) programming interface API application programming interface</p> <p>programming model(s) earth programming model parallel programming model(s)</p> <p>programming paradigm(s) functional programming paradigm</p>
---	--

4.3 Colocados contiguos de las combinaciones especializadas.

PRE +	COMBINACIÓN	+ POST
supports	abstract programming	devices, style, supports, techniques
Common, telephony, XML, easy-to-use, management, high-level, telephone, low-level, API, offer, provide, custom, distributed, standard	application programming	interface(s)
covers, DSPs, symbols	basic programming	constructs, examples, guidelines, ideas, package
threaded-, knowledge of, avoid	C programming	language
teach, recursive	computer programming	consultant, language(s)
demonstrating, permit, types of, principles of, traditional, classical	concurrent programming	assists, concepts, involves, notations, primitives, provides, similar, styles, language(s)
application of, based, use, governing, arise, allow	conventional programming	constructs, language(s), techniques
DSP, FPGA, network	device programming	tools
Set of, covers	distributed programming	activity, models, technology(ies)
use, using	event-driven programming	avoids, model(s)
efficient, developed, identifies, algorithms	dynamic programming	algorithm(s), branch, equations, recurrence, solution, technique
	earth programming	model
use	extreme programming	amplifies, developers, explained, XP
journal of, usual in, concerns, new, typical, parallel, mostly-, concurrent, pure, using, teaching, understand	functional programming	architectures, carries, conference, includes, language(s), paradigm, perspective, practice, research, sense, system, concept(s), logic
	generative programming	addresses, paradigm, techniques
evaluating, think, support, STL-Style, used	generic programming	based, depends, library, techniques, using, GP
measurement-specific, Vee, dataflow-based, G, industry-standard, data-driven	graphical programming	environment(s), language(s)
two-term, excellent, on-line	introductory programming	book, class(es), course(s), sequence, text, textbook
introductory, common, using, advertise	Java programming	books, course, environment(s), jobs, language('s), model, skills

NI, specific, follow, learn	LabVIEW programming	challenge, environment, exercises , experience
application-, high-, G-code, college-, low-	level programming	abstractions , construct, course, include, language
integer , discusses, compact, mixed-integer, derived, non-, integer -, sequential	linear programming	approach, based, formulation , ILP , model, NLP, SLP,
functional , inductive , paradigm , constraint , EPSILON	logic programming	became, concerns, due to, edited, environment, ILP , language(s) , paradigm , studies
	mathematical programming	algorithm, formulation , methods, problem, procedures, solution
object-, aspect -	oriented programming	AOP, COOP, facilities, language(s) , model, paradigms , system(s), technologies
synchronous, full, structured , data, performance, purpose, heavy-weight, multiple, learn , describes	parallel programming	based, concepts , environment, formalism, language, model(s), needs, paradigm(s) , style , system, using, platforms
sequential , via, learning , normal	quadratic programming scheme programming	based, problem, QP , SQP style
no, direct, conventional , corresponding	sequential programming	analogue, language(s) , style , techniques
	software programming	language(s)
represent, concurrent , large	systems programming	environments, including, language(s) , project, usually
end-	user programming	
generic , linear, integer -linear, parallel, quadratic	programming based	methodology, formulation , approach, framework
ASKALON, built-in, integrated, interactive, Java , LabVIEW , logic, multi-, Parallaxis, parallel, pool , target, NET, open, diagram, graphical , different, common, integrated, reasonable, Matlab application-specific, systems	programming environment(s)	supporting, using, allows, CodeWarrior, provide, come, PEs
application(s), visual , PARIS, year, industry-standard	programming interface(s)	API , serve, LAPI , TAPI , well
Ada , C, threaded -C, common, conventional , declarative, Erlang, featured, functional , general, good, graphical , Java , imperative , JavaTM , high-level, logic, major, modern, new, object-oriented , parallel, powerful, proper, pure, scientific, sequential , SR, software, standard, teaching , test, real-time, traditional, true, computer, favorite, first, functional , hardware, normal, purpose, technical, C++, Fortran , full, G, concurrent , existing, Unicon, text-based, batch, full-fledged, mainstream, major, multiple	programming language(s)	utilizing, implementor, initially, developed, directed, beloved, designer, supports, concepts , constructs , analogy, called, counterparts, produced, syntax , terminology, description(s), developed, CFLP, designed, optimization, abstractions , optimized, execution, offering, implementation, characteristics, community, Java , C, objects, point, lisp, compilers, issues, education, feature, object, RMI, description, reflect, manipulate, I/O, aiming, declarations, intended, originate
common, constant, decent, event-driven, EARTH , good, Intel's, Java , JavaBeans , high-level, object-oriented , parallel, SPMD, SPP, straightforward, suitable, synchronous, underlying, MPI's, general, linear, pool , Etherware, based, define, multiple, alternative, geometric, message-passing, purpose, distributed, existing	programming model(s)	based, abstractions , supports, called, hides, described, supported, developed
functional , generative , parallel, SM, logic, OO, different, mixed, presents,	programming paradigm(s)	-specific

<i>shared-addressspace,</i> graphical , <i>object-oriented</i>		
<i>abstract, good, metamorphic,</i> sequential , <i>scheme, parallel</i>	<i>programming style</i>	<i>leads, offers, choice</i>
<i>dynamic, good, abstract, advanced,</i> conventional, generative, generic, <i>describing, including, sequential,</i> <i>recommended, reusable</i>	<i>programming technique(s)</i>	<i>support, impossible,</i> <i>adopted</i>
<i>better, device, invented</i>	<i>programming tool(s)</i>	

5. Familia de *window*. Forma analizada: WINDOWS

1. FRECUENCIA.

KEYWORD	Frec. Telec	Frec. Lacell	Ratio	Término, Chung	KEYNESS	P valor
377. WINDOWS	1.348	1.497	3,41986211	NO	991,6	0
Miembros relacionados:	F Telec	F Gene	Ratio	Término	Keyness	P valor
WINDOW / West	1.228	1.934	2,41147298	NO	531,1	0
WINDOWING	39	4	37,0292674	NO	97,6	0
WINDOWED	7	3	8,86170501	NO	11,1	0,000845
Familia técnica: No estadística y formalmente.						

2. DISTRIBUCIÓN.

Distribución por áreas	Distribución Keyness	Keykeyword	Nº de textos	Porcentaje
9	0	WINDOWS	87	5,26%
<i>Dispersion plot</i>				
Area	Words	Hits	Per 1,000	
082 Esp. Tele	997.727	548	0.55	
3 Telematics	1.205.064	364	0.30	
4 Signal proc.	580.936	128	0.22	
2 Ar. Comp	329.643	67	0.20	
081 Esp. Sign	867.208	112	0.13	
6 Business	373.079	46	0.12	
7 Systems	307.691	26	0.08	
5 Materials	101.241	5	0.05	
1 Electronics	722.823	32	0.04	

3. COLOCACIÓN.

3.1 Colocados.

N	Collocates	TOTAL	LEFT	RIGHT	L5	L4	L3	L2	L1	*	R1	R2	R3	R4	R5
1	THE	605	360	245	88	63	43	39	127	0	1	47	63	59	75
2	AND	420	200	220	29	35	29	38	69	0	25	70	50	43	32
3	A	270	157	113	29	29	13	24	62	0	0	17	31	39	26
4	IN	258	158	100	19	26	20	41	52	0	11	8	43	18	20
5	TO	243	141	102	30	24	33	29	25	0	12	13	29	21	27
6	OF	238	178	60	36	37	28	37	40	0	6	12	8	12	22
7	SERVER	183	36	147	12	6	7	11	0	0	78	30	19	10	10
8	FOR	180	133	47	17	19	14	36	47	0	2	11	5	13	16
9	XP	165	11	154	3	1	0	4	3	0	131	3	10	2	8
10	MICROSOFT	160	140	20	1	3	4	5	127	0	0	7	5	5	3
11	IS	159	72	87	23	18	17	10	4	0	10	23	20	23	11
12	ON	156	116	40	11	13	19	26	47	0	10	8	5	12	5
13	NT	145	24	121	1	9	2	8	4	0	103	4	7	6	1
14	WITH	139	93	46	9	16	12	26	30	0	4	11	10	11	10
15	OR	122	54	68	6	8	6	11	23	0	8	21	16	13	10
16	AS	100	61	39	12	9	11	11	18	0	0	8	8	9	14
17	THAT	99	48	51	10	13	8	7	10	0	6	8	16	12	9
18	SYSTEM	80	23	57	5	5	8	4	1	0	2	11	11	24	9

19	TCP	74	24	50	3	6	8	5	2	0	2	42	3	1	2
20	YOU	73	24	49	7	7	5	5	0	0	0	19	10	11	9
21	FROM	68	47	21	7	11	5	14	10	0	3	6	3	6	3
22	X	65	33	32	1	4	3	3	22	0	12	3	2	10	5
23	BASED	64	13	51	1	4	5	3	0	0	32	12	0	4	3
24	LINUX	58	26	32	3	8	5	8	2	0	1	11	7	6	7
25	CAN	56	13	43	5	5	2	1	0	0	2	4	17	6	14
26	OPERATING	55	19	36	6	6	5	1	1	0	17	11	4	3	1
27	NOT	54	22	32	5	4	9	2	2	0	0	6	12	7	7
28	ARE	53	31	22	7	10	9	2	3	0	1	6	4	6	5
29	BY	53	34	19	8	10	5	6	5	0	5	2	4	0	8
30	UNIX	52	31	21	5	4	7	8	7	0	0	8	4	8	1
31	SOFTWARE	48	18	30	4	2	2	9	1	0	13	3	3	4	7
32	USED	47	27	20	1	6	16	2	2	0	0	3	2	5	10
33	BE	46	24	22	12	4	5	3	0	0	0	6	6	7	3
34	IT	45	17	28	7	6	2	2	0	0	0	10	10	6	2
35	SYSTEMS	45	18	27	5	5	5	2	1	0	5	13	6	1	2

3.2 Colocados significativos.

A) Colocados significativos en el corpus técnico.

Nº	Colocados	Z	T	MI	Nº	Colocados	Z	T	MI
1	XP	241,22	12,81	8,47	45	NETWARE	10,86	2,34	4,43
2	MEDIAPLAYER	73,20	4,11	8,31	46	LINUX	33,34	7,25	4,40
3	REALPLAYER	57,66	3,73	7,90	47	WORKSTATIONS	11,02	2,51	4,27
4	NT	176,35	11,99	7,76	48	MB	10,86	2,50	4,23
5	VXWORKS	29,07	2,22	7,42	49	RUNS	17,77	4,22	4,15
6	DATACENTER	29,07	2,22	7,42	50	IMPLEMENTS	12,52	2,98	4,14
7	PCSPIM	43,60	3,58	7,21	51	CHOSE	8,63	2,10	4,07
8	SWS	26,89	2,22	7,20	52	SCAN	9,91	2,48	4,00
9	MACINTOSH	53,78	4,44	7,20	53	APPLE	9,08	2,29	3,97
10	TIMESTAMPS	25,97	2,22	7,10	54	DEAD	8,18	2,09	3,94
11	MILLENNIUM	36,72	3,14	7,10	55	RUNNING	22,84	5,90	3,90
12	INZERILLO	26,96	2,43	6,94	56	OPENED	7,98	2,08	3,87
13	NTFS	26,30	2,43	6,87	57	PPTP	7,88	2,08	3,84
14	HAMMING	25,69	2,43	6,81	58	WORKSTATION	7,73	2,07	3,79
15	FLAVORS	24,07	2,42	6,62	59	CLIENTS	15,35	4,15	3,78
16	EFS	20,89	2,21	6,48	60	DESKTOP	12,34	3,34	3,77
17	QUICKTIME	20,44	2,21	6,42	61	BUILT	18,81	5,22	3,70
18	DFS	32,79	3,56	6,40	62	DNS	9,67	2,76	3,62
19	MICROSOFT	114,55	12,50	6,39	63	INSTALL	9,51	2,75	3,58
20	SOLARIS	45,82	5,13	6,32	64	DHCP	6,96	2,04	3,54
21	DOS	35,53	4,30	6,10	65	CONSOLE	6,89	2,04	3,51
22	SPIM	17,63	2,20	6,00	66	VERSIONS	11,49	3,41	3,50
23	EDITION	45,20	5,65	6,00	67	DELAYED	7,40	2,23	3,47
24	EXPLORER	32,75	4,17	5,94	68	PCS	10,02	3,02	3,46
25	HACK	16,35	2,20	5,79	69	COMPUTERS	15,68	4,80	3,42
26	PLAYER	28,20	3,80	5,78	70	MS	9,79	3,00	3,41
27	BUNDLE	17,39	2,40	5,71	71	SERVERS	16,36	5,04	3,40
28	PROFESSIONAL	39,38	6,24	5,31	72	MEDIA	16,04	5,09	3,31
29	CE	24,65	4,01	5,24	73	INSTALLED	9,40	2,98	3,31
30	UNIX	40,95	7,00	5,10	74	UPGRADE	6,78	2,19	3,26
31	IGMP	12,17	2,17	4,98	75	ME	8,21	2,68	3,23
32	WINS	11,97	2,16	4,94	76	DRIVER	6,66	2,19	3,22

33	MICROSOFT'S	14,07	2,56	4,92	77	HOSTS	9,77	3,22	3,21
34	HOSTING	16,67	3,06	4,89	78	SUN	7,65	2,52	3,20
35	AVOIDANCE	15,76	2,90	4,89	79	DIRECTORY	14,30	4,72	3,20
36	NATIVE	16,23	3,05	4,82	80	SUPPORTED	11,76	3,88	3,20
37	PACK	11,35	2,16	4,79	81	SHELL	8,52	2,82	3,19
38	HACKING	11,03	2,15	4,72	82	EXTENSIONS	7,06	2,35	3,17
39	REGISTRY	14,34	2,88	4,63	83	DISK	8,35	2,81	3,15
40	JOURNAL	15,09	3,03	4,63	84	CONFIGURE	7,92	2,66	3,15
41	JAPANESE	10,53	2,14	4,59	85	SETTINGS	8,22	2,80	3,11
42	TEMPLATE	14,08	2,87	4,58	86	DIELECTRIC	6,82	2,34	3,09
43	OS	19,17	3,95	4,56	87	PC	12,34	4,23	3,09
44	PLATFORMS	22,94	4,78	4,52	88	PLATFORM	11,08	3,84	3,06

B) Colocados significativos en el corpus general.

Nº	Colocados	Z	T	MI	Nº	Colocados	Z	T	MI
1	MULLIONED	71,01	2,45	9,72	47	TALL	17,51	3,58	4,58
2	CLERESTORY	77,29	2,82	9,55	48	DRUM	10,21	2,14	4,51
3	SASH	84,64	3,60	9,11	49	FLOORS	10,10	2,14	4,48
4	CD	39,82	2,23	8,32	50	ELEGANT	11,06	2,34	4,48
5	MACINTOSH	54,14	3,72	7,72	51	VERSIONS	10,87	2,34	4,44
6	LEXIS	59,64	4,34	7,56	52	PANELS	9,91	2,13	4,43
7	SHUTTERS	31,93	2,44	7,43	53	ELECTRIC	16,41	3,56	4,40
8	STAINED	64,72	4,97	7,41	54	SHAPED	9,67	2,13	4,37
9	DOS	39,76	3,29	7,19	55	LOCK	11,44	2,52	4,37
10	NEXIS	41,61	3,58	7,08	56	USERS	17,02	3,80	4,33
11	SMASHED	42,79	3,71	7,05	57	SERVER	9,43	2,12	4,30
12	PIERCED	31,34	2,81	6,96	58	CLEANING	10,30	2,32	4,30
13	WORDPERFECT	24,19	2,22	6,89	59	WALLS	17,81	4,03	4,29
14	GLAZED	35,57	3,29	6,87	60	USER	14,48	3,29	4,28
15	BLINDS	27,71	2,62	6,80	61	BROKEN	18,09	4,13	4,26
16	ICONS	31,16	2,97	6,78	62	BLANK	9,27	2,12	4,26
17	BARRED	25,06	2,43	6,74	63	SHUT	13,74	3,14	4,26
18	GLAZING	22,33	2,21	6,67	64	ROOF	14,90	3,42	4,25
19	SEMI	24,14	2,42	6,63	65	APPLICATIONS	13,92	3,27	4,18
20	BOARDED	22,01	2,21	6,63	66	GOTHIC	8,96	2,11	4,17
21	MIRRORS	32,35	3,28	6,60	67	OPERATING	13,68	3,27	4,13
22	ARCHED	21,70	2,21	6,59	68	BAY	12,43	2,98	4,12
23	OS	30,47	3,13	6,57	69	IBM	8,66	2,10	4,08
24	MB	20,98	2,21	6,49	70	LIGHTS	13,22	3,25	4,04
25	LOCKING	19,95	2,21	6,35	71	INTERIOR	10,07	2,48	4,04
26	MAC	31,85	3,69	6,22	72	MS	11,10	2,81	3,97
27	TM	18,86	2,21	6,19	73	CEILING	8,19	2,09	3,94
28	CEILINGS	18,47	2,20	6,13	74	HEADED	8,95	2,29	3,93
29	DIFFERS	20,73	2,60	5,99	75	BATHROOM	8,86	2,29	3,91
30	OPENINGS	16,81	2,20	5,87	76	REAR	10,11	2,64	3,88
31	LOCKS	22,67	3,10	5,74	77	X	11,15	2,94	3,84
32	CURTAINS	25,06	3,53	5,65	78	WHEELS	7,86	2,08	3,84
33	ALLOY	15,29	2,19	5,61	79	APPLE	8,60	2,28	3,83
34	ARCHES	16,12	2,40	5,50	80	TINY	11,55	3,08	3,81
35	COMPATIBLE	17,17	2,59	5,46	81	FURNITURE	9,51	2,61	3,73
36	LIT	19,50	3,08	5,32	82	FACING	8,30	2,42	3,56
37	CLEANED	13,26	2,18	5,21	83	BARS	7,59	2,24	3,52
38	CLICK	15,58	2,57	5,20	84	FITTED	8,17	2,41	3,52

39	PC	19,23	3,22	5,15	85	BROKE	7,92	2,40	3,44
40	INTERFACE	12,67	2,17	5,09	86	PLAYER	9,06	2,85	3,34
41	INSTALLED	16,69	2,91	5,04	87	SOUL	6,85	2,20	3,28
42	MATCHING	13,40	2,37	5,00	88	STEEL	6,17	2,00	3,25
43	TOPS	13,32	2,37	4,98	89	FORGET	7,98	2,67	3,16
44	BLOWN	11,68	2,16	4,87	90	PROGRAMS	7,50	2,51	3,16
45	LOTUS	11,43	2,16	4,81	91	FACES	6,10	2,15	3,01
46	CIRCULAR	12,38	2,36	4,78					

3.3 Definición del nodo en el diccionario.

1. *Window* se encuentra registrada en la *General Service List* (West, 1953) pero no se ejemplifica su uso.
2. Definición de *window* en el *Advanced Learner's Dictionary* de Cambridge (2003):

Window (noun):

1. (glass) A space usually filled with glass in the wall of a building or in a vehicle, to allow light and air in and to allow people inside the building to see out.
2. (computer) A separate area on a computer screen which shows information and which you can move around.
3. (opportunity) A period when there is an opportunity to do something.

3. Definición de *window* en el diccionario especializado Webster:

Windows (noun), plural of **window**,

Specialty Definition:

In electrical engineering:

1. An aperture produced in a silicon-oxide layer covering a semiconductor surface; used to define a diffusion mask or to enable electrical contact to be made to the surface.
2. The ordered set of *W* consecutive packet send sequence numbers *P(S)* of the data packets authorized to cross the DTE/DCE interface of a logical channel used for a virtual call or permanent virtual circuit and for each direction of transmission.
3. Glass, plastic or other transparent material used to cover collector-absorber plate so that the solar energy is "trapped" by the greenhouse effect.

In post & telecom:

1. In X. 25: ordered set of window consecutive packet send sequence numbers of the data packets authorized to cross the interface.

In computing:

A portion of a computer screen, usually rectangular in shape, showing the output of and allowing input for one of a number of simultaneously running computer processes. The size and position can normally be adjusted, for example covering the whole screen, or half the screen, allowing another non-overlapping window using the other half, etc. Windows are a popular feature (or widget) in several graphical user interfaces. The X Window System and Microsoft Windows are named after this feature.

Microsoft Windows is a range of operating environments for personal computers. The range was first introduced by Microsoft in 1985 and eventually came to dominate the world personal computer market. All recent versions of Windows also function as a fully-fledged operating system.

3.4 Colocados contiguos.

a) Colocado + windows.		
Total colocados seleccionados: 15	Frecuencia máxima: 127	Frecuencia mínima: 4
<ul style="list-style-type: none"> - Sustantivos, total 4: <i>Microsoft</i> 127, <i>hosts</i> 13, <i>Realplayer</i> 4, <i>bit</i> 4. - Participios verbales, total 3: <i>running</i> 14, <i>using</i> 7, <i>hosting</i> 4. - Formas personales de verbos, total 2: <i>use</i> 10, <i>configure</i> 4. - Adjetivos, total 4: <i>dielectric</i> 5, <i>Standard</i> 4, <i>native</i> 4, <i>earlier</i> 4. - Formas abreviadas, total 2: <i>X</i> 22, <i>MS</i> 7. 		

b) <i>Windows</i> + colocado.		
Total colocados seleccionados: 19	Frecuencia máxima: 131	Frecuencia mínima: 5
<ul style="list-style-type: none"> - Sustantivos, total 11: <i>server</i> 78, <i>media</i> 15, <i>mediaplayer</i> 15, <i>software</i> 13, <i>explorer</i> 12, <i>journal</i> 10, <i>applications</i> 7, <i>program</i> 7, <i>version</i> 7, <i>environment</i> 6, <i>platform</i> 5. - Participios verbales, total 3: <i>based</i> 32, <i>operating</i> 17, <i>supported</i> 5. - Formas abreviadas, total 5: <i>XP</i> 131, <i>NT</i> 103, <i>X</i> 12, <i>R</i> 5. 		

3.5 Identificación de colocados contiguos especializados en las líneas de concordancia.

*rture coupled microstrip antennas. PCAAD 3.0 is a **Windows-based** general purpose antenna analysis ment in general. "The reason [Cisco] goes with a **Windows-based** operating system on its [voice-ove useful to IT staff looking to support Linux users in a **Windows-based** network who need to access Win tate 'initial' The RolEnact Interface. RolEnact is a **Windows-based** application written in Enact, a hy rt time or visit client sites regularly, if you rely on a **Windows-based** laptop to connect to multiple netw es out of ten, someone surfing the Web is using a **Windows-based** machine. This could prove frustrati tem. The PC houses Cohu's CAMS 2.0 software, a **windows based** application that will manage and c C. · The PC houses Cohu's CAMS 2.0 software, a **windows-based** application that will manage and c ncurrently run Linux-based management tools and **Windows-based** desktop applications. "Once [Wi Solaris Unix server and can backup files from any **Windows-based** PC or laptop, as well as Linux-ba and spam-fighting issues. While part IV looked at **Windows-based** solutions, along with appliances a interface (called MICE), and adapted a commercial **Windows based** DUA/UA interface for their own us o 14 in.) and that are all configured with a common **Windows-based** configuration software package t l office system was being upgraded to a graphical, **windows-based** system. Not a lot of planning or te ty of interfaces were installed and tested, including windows based ones and dial-up modem access vi software that allows Linux clients to view Microsoft **Windows-based** plug-ins embedded in Web pages. rdware failure or user error. It consists of a Web- or **Windows-based** administrative console, the Windo k adds true AMX-awareness to a variety of popular **Windows based** debuggers. Development Informat servers more stable and scalable than its previous **Windows-based** ECLIPS products. Competing IP generic Windows browser access, and proprietary **Windows-based** software. Generic browser access ation File System. The capability for users to run **Windows-based** Lotus Notes messaging clients on uffer between a Linux client's Web browser and the **Windows-based** plug-in, such as embedded MS Of and what's more, they are "100% Pure G". For the **Windows-based** LabVIEW developer, this means t r and Nortel's Business Communication Server use **Windows-based** operating systems that run the co rior to the current de-facto standard X windows. - X **windows-based** tools such as XTV [AbFe91], shX [ssion Line Calculator. Txline is a free easy-to-use, **Windows-based**, interactive transmission line calc*

Concordancias de *Windows-based*.

a) Colocado + *windows*:

<i>bit Windows</i>	native Windows
dielectric Windows	Realplayer Windows
hosting Windows	running Windows
hosts Windows	<i>standard Windows</i>
Microsoft Windows	<i>X Windows</i>
MS Windows	

b) *Windows* + colocado:

<i>Windows application(s)</i>	<i>Windows operating</i>
<i>Windows based</i>	<i>Windows platform(s)</i>
<i>Windows CE</i>	<i>Windows program(s)</i>
<i>Windows environment</i>	<i>Windows server(s)</i>
<i>Windows explorer</i>	<i>Windows socket(s)</i>
<i>Windows journal</i>	<i>Windows software</i>
<i>Windows media</i>	<i>Windows supported</i>
<i>Windows mediaplayer</i>	<i>Windows version</i>
<i>Windows NT</i>	<i>Windows XP</i>

4. AGRUPACIONES LÉXICAS O CLUSTERS.

4.1 Distribución de clusters.

Combinaciones	Volumen	Frecuencia máxima
2-cluster	125	137
3-cluster	138	9
4-cluster	86	13
5-cluster	36	13
6-cluster	28	9

4.2 Inclusión de combinaciones especializadas en unidades más amplias.

<p>Microsoft Windows <i>running Microsoft Windows</i> <i>Microsoft Windows XP</i> <i>Microsoft Windows XP computer</i> <i>Microsoft Windows (n°) DHCP</i> <i>EFS in Microsoft Windows XP</i> <i>security template for Microsoft Windows</i> <i>template for Microsoft Windows XP</i> <i>Microsoft Windows (n°) DHCP server</i> <i>security template for Microsoft Windows XP</i></p> <p>Windows NT <i>Microsoft Windows NT</i> <i>Windows NT server(s)</i> <i>Windows NT systems</i> <i>Windows NT domain</i></p> <p>Windows operating <i>Windows operating system(s)</i></p>	<p>Windows software <i>Windows software package</i></p> <p>Windows supported <i>Windows supported target</i> <i>Windows supported target processors</i></p> <p>Windows XP <i>use Windows XP</i> <i>Windows XP computer</i> <i>Windows XP workstations</i> <i>Windows XP professional computer(s)</i> <i>Microsoft Windows XP computer</i> <i>EFS in Microsoft Windows XP</i> <i>security template for Microsoft Windows XP</i></p> <p>Windows CE <i>Windows CE operating</i> <i>Windows CE operating system</i></p>
--	---

4.3 Colocados contiguos de las combinaciones especializadas.

PRE +	COMBINACIÓN	+ POST
<i>deep</i>	<i>dielectric Windows</i>	<i>using</i>
<i>enterprise</i>	<i>hosting Windows</i>	<i>bundle</i>
<i>development</i>	<i>hosts Windows</i>	Linux, supported
<i>running, using, including, view, dozen, software, well-designed, IP-forwarding, Micromass, system, built-into, hack, dialup, batch, use, group, interactive, service, users</i>	Microsoft Windows	<i>2000, 2003, 95, -based, boxed, CE, clusters, current, environment, GUI, machine, NT, operating, platforms, provide, security, ships, style, supplies, update, XP, Media, software, server, TM, current</i>
<i>preparing</i>	MS Windows	<i>application, interfaces, de facto</i>
<i>including</i>	native Windows	<i>application, support</i>
<i>server(s), clients, memory, Celeron</i>	running Windows	<i>2000, 95, 98, NT, IBM, program</i>
<i>X</i>	<i>standard Windows</i>	desktop, functions, installation, NT
<i>protocol, standard, re-scanning, compatible, framestore, GUI, independent</i>	<i>X Windows</i>	<i>-based, servers, workstations, motif, compatible, environment, resolutions, software</i>
<i>includes, Linux, configure, install</i>	<i>Windows 2000</i>	<i>-based, /XP</i>
<i>interactive, MS, native, simple</i>	<i>Windows application(s)</i>	<i>developed, development</i>
<i>commercial, common, including, Microsoft, popular, previous, proprietary, run, use, easy-to-use, X</i>	<i>Windows-based</i>	<i>general, operating, network, application, laptop, machine, desktop, PC, solutions, DUA, system, configuration, plug-in(s), tools, administrative, debuggers, ECLIPS, software, Lotus, LabVIEW</i>
Microsoft	<i>Windows CE</i>	<i>leadership, platform, strategy, operating, market, -powered</i>
<i>X-, Microsoft</i>	<i>Windows environment</i>	
	<i>Windows explorer</i>	<i>loading, program,</i>

<i>automatic, familiar</i>		<i>application, tool</i>
<i>opens,</i>	<i>Windows journal</i>	<i>allows</i>
<i>Microsoft's, need, supports</i>	<i>Windows media</i>	<i>center, player, software</i>
	<i>Windows media</i>	<i>required, version, fluctuated, setup</i>
<i>aging, chose, Novell-connected, Microsoft, non-, redundant, remote, seconds, standard, TCP/IP, using, running</i>	<i>Windows NT</i>	<i>server(s), device, domain(s), native, systems, synchronizes, workstation, based, embedded, environments, reports, administration, provides, operating, -based, /2000</i>
<i>DOS/, earlier, familiar, Microsoft, non-</i>	<i>Windows operating</i>	<i>system(s)</i>
<i>Microsoft</i>	<i>Windows platform(s)</i>	<i>does</i>
<i>32-bit, every, running, containing, most, co-operating</i>	<i>Windows program(s)</i>	<i>package, running, tried</i>
<i>combining, implementing, improvements, Microsoft®, multiple, support, X</i>	<i>Windows server(s)</i>	<i>2003, platforms, market</i>
<i>using</i>	<i>Windows socket(s)</i>	<i>setsockopt, Winsock, function</i>
<i>Microsoft's, X-</i>	<i>Windows software</i>	<i>package, codes, new, meant, runs, running</i>
	<i>Windows supported</i>	<i>target</i>
<i>earlier, switch</i>	<i>Windows version</i>	
<i>department's, deploy, install, Microsoft, platform, use, garden-variety</i>	<i>Windows XP</i>	<i>Solaris, professional, service, workstations, SP, tablet, built-in, holds, backup, computer(s), applies, domain, single, uniquely, identifies, contains, media, embedded, SP1, Japanese, operating, home</i>

6. Familia de bus. Forma analizada: BUS

1. FRECUENCIA.

KEYWORD	Frec. Telec	Frec. Lacell	Ratio	Término, Chung	KEYNESS	P valor
980. BUS / West	711	1.197	2,25587979	NO	269,7	0
Miembros relacionados:	F Telec	F Gene	Ratio	Término	Keyness	P valor
BUSSES	16	1	60,7659772	ESPECIFICO	43	0
BUSWAY	12	3	15,1914943	NO	24	0,000001
BUSWAYS	9	1	34,1808622	NO	22,2	0,000002
BUSU	4	0	inf/esp	inf/esp	12,5	0,000397
Familia técnica: No, estadística y formalmente.						

2. DISTRIBUCIÓN.

Distribución por áreas	Distribución Keyness	Keykeyword	Nº de textos	Porcentaje
9	0	BUS	48	2,9%
<i>Dispersion plot</i>				
Area	Words	Hits	Per 1,000	
2 Ar. Comp	329.643	267	0.81	
1 Electronics	722.823	127	0.18	
082 Esp. Tele	997.727	90	0.09	
3 Telematics	1.205.064	84	0.07	
4 Signal proc.	580.936	56	0.10	
7 Systems	307.691	43	0.14	
081 Esp. Sign	867.208	32	0.04	
6 Business	373.079	17	0.05	
5 Materials	101.241	0	0	

3. COLOCACIÓN.

3.1 Colocados.

N	Collocates	TOTAL	LEFT	RIGHT	L5	L4	L3	L2	L1	*	R1	R2	R3	R4	R5
1	THE	631	394	237	49	46	63	109	127	0	0	73	57	47	60
2	A	275	175	100	19	22	27	51	56	0	0	35	25	18	22
3	AND	198	78	120	19	16	18	14	11	0	20	39	19	24	18
4	TO	161	75	86	10	17	30	16	2	0	15	8	16	24	23
5	IS	141	40	101	11	12	13	3	1	0	30	25	19	17	10
6	OF	136	83	53	17	11	15	26	14	0	9	7	6	14	17
7	IN	95	37	58	11	8	8	7	3	0	3	23	8	17	7
8	ON	78	63	15	7	14	14	27	1	0	1	2	4	5	3
9	FOR	69	30	39	12	9	5	3	1	0	6	10	8	7	8
10	DATA	59	48	11	5	9	4	0	30	0	0	2	1	5	3
11	WITH	53	29	24	3	5	9	11	1	0	2	8	5	3	6
12	ARE	52	18	34	10	5	3	0	0	0	2	10	8	9	5
13	SYSTEM	52	34	18	4	1	3	17	9	0	4	2	5	6	1
14	AS	48	29	19	8	7	7	5	2	0	2	6	3	6	2
15	OR	44	21	23	1	7	4	4	5	0	9	7	1	5	1
16	THIS	43	12	31	4	5	0	1	2	0	0	13	6	7	5
17	THAT	42	19	23	6	3	6	2	2	0	8	9	3	2	1

18	AN	39	25	14	3	11	5	6	0	0	0	4	3	3	4
19	ADDRESS	36	21	15	1	3	3	0	14	0	1	3	6	3	2
20	ARCHITECTURE	36	9	27	1	2	4	1	1	0	19	2	1	3	2
21	BE	35	16	19	10	5	1	0	0	0	0	3	5	6	5
22	BY	34	19	15	4	4	2	7	2	0	3	5	2	2	3
23	BIT	33	29	4	2	1	12	12	2	0	0	1	1	0	2
24	CAN	33	7	26	2	0	3	0	2	0	4	3	8	8	3
25	POWER	33	28	5	3	1	5	1	18	0	1	1	1	2	0
26	PCI	32	29	3	0	0	1	1	27	0	1	0	0	1	1
27	AT	29	8	21	3	0	2	3	0	0	7	5	6	3	0
28	WHICH	28	7	21	2	2	1	1	1	0	4	7	2	5	3
29	IT	27	5	22	0	3	2	0	0	0	0	9	8	3	2
30	TWO	25	17	8	5	7	1	2	2	0	0	3	0	2	3
31	ARCHITECTURES	24	5	19	4	1	0	0	0	0	17	1	0	0	1
32	FROM	24	18	6	1	5	4	8	0	0	2	1	2	1	0
33	INTERFACE	23	10	13	1	4	2	0	3	0	11	1	1	0	0
34	HAVE	22	2	20	2	0	0	0	0	0	0	1	2	13	4
35	MEMORY	22	15	7	3	2	1	3	6	0	0	0	4	1	2

3.2 Colocados significativos.

A) Colocados significativos en el corpus técnico.

Nº	Colocados	Z	T	MI	Nº	Colocados	Z	T	MI
1	PCI	84,36	5,63	7,81	23	TRANSFERS	15,44	2,74	4,99
2	BUS'S	74,78	2,45	9,87	24	HOT	15,25	2,89	4,80
3	LVDS	62,47	3,31	8,48	25	BIDIRECTIONAL	15,16	2,39	5,33
4	BACKPLANES	57,98	3,73	7,92	26	PLUG	14,63	2,38	5,23
5	ISA	50,04	2,99	8,13	27	MINIMIZE	13,42	2,86	4,46
6	SETTLING	44,46	2,82	7,96	28	CPU	13,42	3,38	3,98
7	DMA	41,06	2,98	7,56	29	TOPOLOGIES	12,73	2,54	4,65
8	BUSES	34,56	3,43	6,67	30	WIRES	12,54	2,70	4,43
9	VME	32,80	2,63	7,28	31	TI	11,96	2,69	4,31
10	GROUPLLET	30,49	3,28	6,43	32	INTERRUPT	11,18	2,34	4,51
11	LEC	29,56	2,80	6,80	33	CYCLE	9,65	2,88	3,49
12	USB	26,86	3,12	6,21	34	CONVERTERS	9,63	2,13	4,36
13	TRANSCEIVER	26,53	3,41	5,92	35	DRIVE	9,28	2,61	3,66
14	DCMM	26,24	2,43	6,87	36	INTERMEDIATE	8,55	2,43	3,63
15	CONVERTER	24,89	3,90	5,35	37	ATTACHED	8,37	2,56	3,41
16	LES	23,66	2,42	6,57	38	COMPATIBLE	8,37	2,42	3,58
17	BACKPLANE	21,28	2,94	5,71	39	INSTRUCTION	8,14	2,55	3,35
18	BATTERY	19,22	3,49	4,93	40	EIGHT	7,48	2,23	3,49
19	STAR	18,79	2,93	5,36	41	FLOATING	7,42	2,06	3,69
20	SERIAL	17,16	3,80	4,35	42	WIRE	6,83	2,34	3,10
21	NOTIFICATION	17,08	2,75	5,26	43	DUAL	6,13	2,15	3,02
22	UNIVERSAL	16,30	3,19	4,71					

B) Colocados significativos en el corpus general.

Nº	Colocados	Z	T	MI
1	TIMETABLES	30,73	2,22	7,58
2	DEPOT	43,50	3,30	7,44
3	TAXI	40,72	3,96	6,72
4	FARES	30,19	2,97	6,69
5	DEREGULATION	22,37	2,21	6,67
6	FARE	37,26	3,70	6,66

7	STOPS	42,34	4,53	6,45
8	CONDUCTOR	27,63	2,97	6,44
9	UNLIMITED	26,81	2,96	6,36
10	TROLLEY	18,78	2,21	6,18
11	ROUTES	34,35	4,07	6,16
12	LANES	19,96	2,41	6,10
13	MINI	22,09	2,78	5,98
14	CANAL	31,82	4,06	5,94
15	DRIVERS	34,80	4,51	5,90
16	CASUALTIES	18,25	2,59	5,63
17	PASSENGERS	27,97	4,04	5,58
18	PASSENGER	23,46	3,52	5,47
19	RIDE	31,77	4,79	5,46
20	BUSES	19,78	3,09	5,36
21	RAIL	26,46	4,14	5,35
22	SHELTER	15,92	2,58	5,25
23	ROUTE	27,07	4,46	5,20
24	AMSTERDAM	13,95	2,38	5,11
25	WISE	18,72	3,22	5,08
26	CATCH	25,10	4,34	5,06
27	TICKET	18,23	3,21	5,01
28	OPERATORS	12,72	2,36	4,86
29	ACCIDENTS	11,81	2,35	4,66
30	LOTHIAN	9,86	2,13	4,42
31	NEARBY	12,18	2,69	4,36
32	COACH	10,27	2,64	3,92
33	BOAT	10,08	2,64	3,87
34	LANE	8,29	2,27	3,74
35	SCHEDULE	7,21	2,05	3,62
36	TRAVELLING	6,96	2,04	3,54
37	TICKETS	6,87	2,04	3,51

3.3 Definición del nodo en el diccionario.

1. Significado de *bus* en la *General Service List* (West, 1953):

Bus (noun): *Go by bus; get a bus; catch a bus.*

2. Definición de *bus* en el *Advanced Learner's Dictionary* de Cambridge (2003):

Bus (noun): *A large vehicle in which people are driven from one place to another.*

3. Definición de *bus* en el diccionario especializado Webster:

Bus (noun): *A vehicle carrying many passengers; used for public transport; "he always rode the bus to work".*

Specialty Definition:

Bus (noun): *One of the sets of conductors (wires, PCB tracks or connections in an integrated circuit) connecting the various functional units in a computer.*

In computer architecture:

A bus is a subsystem that transfers data or power between computer components inside a computer or between computers. Unlike a point-to-point connection, a bus can logically connect several peripherals over the same set of wires.

Early computer buses were literally parallel electrical buses with multiple connections, but the term is now used for any physical arrangement that provides the same logical functionality as a parallel electrical bus. Modern computer buses can use both parallel and bit-serial connections, and can be wired in either a multidrop (electrical parallel) or daisy chain topology, or connected by switched hubs, as in the case of USB.

3.4 Colocados contiguos.

a) Colocado + bus.		
Total colocados seleccionados: 40	Frecuencia máxima: 30	Frecuencia mínima: 1
<ul style="list-style-type: none"> - Sustantivos, total 20: <i>Data 30, power 18, address 14, grouplet 7, memory 6, speed 6, control 5, supply 5, wire 5, host 5, side 3, server 3, error 3, swing 3, output 2, input 2, channel 2, software 1, protocol 1, message 1.</i> - Adjetivos, total 13: <i>common 15, serial 12, new 6, intermediate 6, conventional 6, hot 5, dual 5, traditional 3, same 2, ideal 2, different 2, local 2, bidirectional 2.</i> - Participios verbales que actúan como adjetivos, total 2: <i>proposed 6, shared 11.</i> - Formas personales de verbos, total 1: <i>minimize 8.</i> - Siglas, total 6: <i>PCI 27, DC 10, VME 7, PC 4, ISA 4, CMOS 3.</i> 		
b) Bus + colocado.		
Total colocados seleccionados: 31	Frecuencia máxima: 19	Frecuencia mínima: 1
<ul style="list-style-type: none"> - Sustantivos, total 27: <i>Architecture 19, architectures 17, converter 13, transceiver 12, interface 11, topology 11, voltage 8, systems 8, name 8, cycle 5, states 5, plug 5, controller 4, converters 4, process 4, wires 4, error 3, network 3, networks 3, technologies 3, bandwidth 3, transfer 3, lines 3, service 2, protocols 2, protocol 1, results 1.</i> - Adjetivos, total 1: <i>compatible 3.</i> - Participios verbales que actúan como adjetivos, total 1: <i>settling 8.</i> - Siglas, total 2: <i>LVDS 5, USB 4.</i> 		

3.5 Identificación de colocados contiguos especializados en las líneas de concordancia.

and the efficiency of the proposed algorithm. Address Bus Encoding Techniques for System-Level Power data mode, the data bus is signal F_P and the address bus is T_DIR. The second one is the output of a 5-bits bus is 16 bits wide and is nonmultiplexed. The address bus is 24 bits wide. Note that A0, the least significant devices that there is a valid address on the address bus and the processor is synchronized to enable. This memory protection o 16-bit data bus o 24-bit address bus which makes 16-MB physical addressable memory for the marks generation are the following: 1. Address bus is pointing at address "00000". So the module is to the accumulator. The control unit sets the address bus to location 112 and puts the value of the signal L_E indicates read or run mode, the address bus is signal V_EST, the measured input variable. The Note that A0, the least significant bit of the address bus is not output. This bit is used internally in conjunction (program, X and Y data) has its own data and address bus, and all of these connect to the outside world via buses can generally be divided in three types: address bus, data bus, or control bus. System Buses. System an 8-bit data bus, denoted H0-H7, a 3-bit address bus, denoted as HA0-HA2, and a control bus with four micro-instruction decoder through the internal address bus; if it is a non-branching micro-code, the sequencer

Concordancias de address bus.

a) Colocado + bus:

<i>address bus</i>	<i>grouplet bus</i>	<i>PCI bus</i>
<i>control bus</i>	<i>memory bus</i>	<i>PC bus</i>
<i>CMOS bus</i>	<i>supply bus</i>	<i>VME bus</i>
<i>channel bus</i>	<i>host bus</i>	<i>ISA bus</i>
<i>data bus</i>	<i>swing bus</i>	
<i>DC bus</i>	<i>serial bus</i>	

b) Bus + colocado:

<i>bus architecture(s)</i>	<i>bus state(s)</i>	<i>bus network(s)</i>
<i>bus converter</i>	<i>bus control/controller</i>	<i>bus protocol(s)</i>
<i>bus transceiver</i>	<i>bus LVDS</i>	<i>bus technology/technologies</i>
<i>bus topology</i>	<i>bus cycle</i>	<i>bus transfer(s)</i>
<i>bus system(s)</i>	<i>bus interface(s)</i>	

4. AGRUPACIONES LÉXICAS O CLUSTERS.

4.1 Distribución de clusters.

Combinaciones	Volumen	Frecuencia máxima
2-cluster	110	131
3-cluster	89	19
4-cluster	41	11
5-cluster	20	8
6-cluster	17	8

4.2 Inclusión de combinaciones especializadas en unidades más amplias.

<p>CMOS bus conventional CMOS bus</p> <p>data bus bit data bus</p> <p>bus converter DC bus converter DC bus converters</p> <p>memory bus single memory bus a single memory bus</p> <p>swing bus low swing bus</p> <p>serial bus universal serial bus backplane serial bus</p>	<p>bus USB serial bus USB universal serial bus USB</p> <p>bus transceiver speed bus transceiver universal bus transceiver wire bus transceiver</p> <p>bus architecture swing bus architecture low swing bus architecture conventional bus architecture the conventional bus architecture</p>
---	---

4.3 Colocados contiguos de las combinaciones especializadas.

PRE +	COMBINACIÓN	+ POST
(n°)-bit, internal	address bus	encoding techniques, makes, is pointing, location
	control bus	interface, performs
Conventional, voltage	CMOS bus	architecture(s)
I/O, individual	channel bus	-minus outputs
drive, (n°)-bit, physical, bidirectional, high-speed, serial, complete, use, host,	data bus	width, is (n°) bits wide
Full-bridge, 220W, use, output	DC bus	converter(s)
	grouplet bus	should support,
single, external, each	memory bus	interfaces
Second, additional, standard	supply bus	
fibre channel, proprietary, include, invented, host,	host bus	adapter(s), interface
Down low, up low, up-down low	swing bus	DLB, ULB, UDLB, architecture,
	serial bus	headset, USB, connection, ports
bypass, implements, traditional, standard, basic, a (n°)-bit, include, Peripheral Component Interconnect	PCI bus	systems, card runs, interface, controllers, traffic, add-in card, specification
original, simply	PC bus	uses, ran
Standard, (v) use	VME bus	slave, master
PC, older	ISA bus	
conventional, high-speed, new, up low swing, up-down swing, intermediate, CMOS, proposed, single-, ideal, low power, standard, different, real, hierarchical, on-chip, shared, traditional, PI-	bus architecture(s)	proposed, described, compared, shown, considered ULB, UDLB, IBA, results
DC, intermediate, halfbridge	bus converter	need
Universal, high-speed, three-wire	bus transceiver	optimizes, transmits, provides

<i>local, to power, PCI-X, microprocessor, PCI, control, host, memory</i>	<i>bus interface(s)</i>	<i>chip, took, used</i>
<i>Ethernet, to use, coaxial, local</i>	<i>bus topology</i>	<i>LAN</i>
<i>New, shared-, trolley-, telematic, modern low voltage, PCI</i>	<i>bus system(s)</i>	<i>depends on, to be accepted</i>
<i>Sequence of, is called</i>	<i>bus state(s)</i>	
	<i>bus LVDS</i>	<i>capabilities, is used</i>
<i>a special sequence of, the beginning of, uses, is called</i>	<i>bus cycle</i>	<i>consists of</i>
<i>Use, PCI, CPU</i>	<i>bus control/controller</i>	<i>signals, coordinates</i>
<i>Larger, proprietary</i>	<i>bus network(s)</i>	<i>work</i>
<i>types of</i>	<i>bus transfer(s)</i>	<i>Takes place, requests</i>
<i>communications</i>	<i>bus protocol(s)</i>	
<i>new, shared-, traditional, current</i>	<i>bus technology/ies</i>	<i>supports</i>

7. Familia de map. Forma analizada: MAPPING

1. FRECUENCIA.

KEYWORD	Frec. Telec	Frec. Lacell	Ratio	Término, Chung	KEYNESS	P valor
278. MAPPING	631	102	23,4946885	NO	1.435,40	0
Miembros relacionados:	F Telec	F Gene	Ratio	Término	Keyness	P valor
MAPPED	240	56	16,276601	NO	491,8	0
MAPPINGS	148	4	140,521322	ESPECIFICO	429,1	0
MAP / West	502	766	2,48894587	NO	230,1	0
ROADMAP	35	4	33,2313938	NO	85,9	0
IMAP	27	0	inf/esp	inf/esp	84,7	0
MAPS / West	177	274	2,45337089	NO	79	0
MAPPER	14	1	53,1702301	ESPECIFICO	37	0
MAPPOINT	6	0	inf/esp	inf/esp	18,8	0,000014
NMAP	6	0	inf/esp	inf/esp	18,8	0,000014
MAPPERS	7	1	26,585115	NO	16,4	0,000051
REMAPPING	7	1	26,585115	NO	16,4	0,000051
REMAP	4	0	inf/esp	inf/esp	12,5	0,000397
SECMAP	4	0	inf/esp	inf/esp	12,5	0,000397
TRAILMAPPER	4	0	inf/esp	inf/esp	12,5	0,000397

Familia técnica: No, estadística y formalmente.

2. DISTRIBUCIÓN.

Distribución por áreas	Distribución Keyness	Keykeyword	Nº de textos	Porcentaje
9	2	MAPPING	62	3,75%
<i>Dispersion plot</i>				
Area	Words	Hits	Per 1,000	
2 Ar. Comp	329.643	79	0.24	
7 Systems	307.691	67	0.22	
082 Esp. Tele	997.727	157	0.16	
1 Electronics	722.823	108	0.15	
4 Signal proc.	580.936	60	0.10	
3 Telematics	1.205.064	110	0.09	
5 Materials	101.241	8	0.08	
081 Esp. Sign	867.208	37	0.04	
6 Business	373.079	5	0.01	

3. COLOCACIÓN.

3.1 Colocados.

N	Collocates	TOTAL	LEFT	RIGHT	L5	L4	L3	L2	L1	*	R1	R2	R3	R4	R5
1	THE	481	271	210	44	56	21	48	102	0	24	65	29	40	52
2	A	220	146	74	19	17	18	39	53	0	9	20	11	15	19
3	OF	211	85	126	22	12	13	18	20	0	80	7	6	17	16
4	TO	203	84	119	15	12	28	25	4	0	20	9	25	39	26
5	AND	186	87	99	11	13	16	28	19	0	24	12	18	29	16
6	IS	126	58	68	13	17	16	12	0	0	27	11	8	10	12
7	FOR	105	59	46	7	8	9	13	22	0	13	11	8	6	8

8	IN	95	38	57	17	6	8	6	1	0	11	10	10	15	11
9	THIS	63	39	24	8	7	2	4	18	0	3	7	4	5	5
10	THAT	50	25	25	8	2	3	10	2	0	6	6	4	3	6
11	BY	46	35	11	4	6	5	7	13	0	0	2	3	5	1
12	ARE	44	7	37	4	2	0	1	0	0	4	5	6	20	2
13	BE	43	21	22	11	7	3	0	0	0	0	8	2	4	8
14	AN	41	20	21	2	5	0	13	0	0	3	6	4	3	5
15	ON	39	21	18	4	4	3	8	2	0	1	4	4	3	6
16	BETWEEN	38	4	34	3	1	0	0	0	0	30	2	1	0	1
17	AS	37	23	14	4	8	3	5	3	0	2	2	1	4	5
18	WITH	37	21	16	6	2	8	5	0	0	2	3	6	3	2
19	FROM	33	7	26	2	1	4	0	0	0	17	5	2	0	2
20	ADDRESS	30	24	6	2	1	1	2	18	0	0	2	1	2	1
21	STATE	30	9	21	1	2	3	2	1	0	1	16	2	0	2
22	WE	28	15	13	3	6	6	0	0	0	2	3	1	2	5
23	INTO	27	3	24	1	1	0	1	0	0	0	3	4	14	3
24	CAN	26	13	13	5	4	4	0	0	0	8	2	1	1	1
25	ONE	26	17	9	1	2	7	1	6	0	1	3	2	2	1
26	ONTO	26	3	23	1	0	0	2	0	0	3	2	5	7	6
27	VARIABLES	26	10	16	4	0	2	2	2	0	0	0	15	1	0
28	NEXT	25	11	14	2	3	2	0	4	0	14	0	0	0	0
29	NETWORK	24	15	9	4	5	2	3	1	0	1	1	2	2	3
30	OR	24	14	10	2	1	1	4	6	0	2	1	3	3	1
31	DATA	22	6	16	1	3	1	0	1	0	2	4	6	2	2
32	BASED	21	13	8	2	2	6	2	1	0	0	3	3	1	1
33	IP	21	9	12	3	2	0	3	1	0	0	4	1	2	5
34	PROBLEM	19	5	14	0	1	3	1	0	0	10	1	1	1	1
35	WHICH	18	6	12	1	2	1	2	0	0	5	2	1	1	3

3.2 Colocados significativos.

A) Colocados significativos en el corpus técnico.

Nº	Colocados	Z	T	MI
1	SNMP	6,78	2,03	3,47
2	SCHEMA	10,80	2,50	4,22
3	RE	6,28	2,16	3,08
4	PCA	53,00	3,15	8,14
5	ONTO	32,23	4,98	5,39
6	MAP	7,17	2,22	3,39
7	INCOMING	6,99	2,21	3,33
8	GNU	28,41	3,27	6,24
9	FEASIBLE	8,95	2,11	4,17
10	EXPLICIT	7,00	2,05	3,55
11	EDGES	14,83	3,03	4,58
12	DIFFSERV	14,10	2,18	5,38
13	DEN	22,23	2,78	6,00
14	DEFINES	11,68	2,96	3,96
15	ADMINISTRATOR	7,04	2,05	3,56
16	AD	8,26	2,27	3,73
17	ACTORS	9,30	2,12	4,27

B) Colocados significativos en el corpus general.

Nº	Colocados	Z	T	MI
1	ADDRESS	14,40	2,18	5,44
2	FIG	18,70	2,20	6,17
3	NATIONAL	7,45	2,06	3,70
4	STORE	18,82	2,21	6,19

3.3 Definición del nodo en el diccionario.

1. Significado de *map* en la *General Service List* (West, 1953):

Map (noun): *A map of England.*

Map (verb): *Map the country.*

2. El *Advanced Learner's Dictionary* de Cambridge (2003) no registra la definición de *mapping*. El lema más cercano a *mapping* en este diccionario es *map*:

Map (noun):

1. *A drawing of the Earth's surface, or part of that surface, showing the shape and position of different countries, political borders, natural features such as rivers and mountains, and artificial features such as roads and buildings.*
2. *A drawing that gives you a particular type of information about a particular area.*
3. *Something which shows the position of stars in the sky or the features on the surface of planets.*

Map (verb): *to represent an area of land in the form of a map: Parts of the mountainous region in the north of the country have still not been mapped.*

3. Definición de *mapping* en el diccionario especializado Webster:

Mapping (noun): *A function such that for every element of one set there is a unique element of another set.*

Specialty Definition:

In computing science: In computer science, a mapping is the relationship between a key and its value in an associative array.

In formal logic, the term "mapping" is sometimes used for a functional predicate, whereas a function is a model of such a predicate in set theory.

3.4 Colocados contiguos.

a) Colocado + <i>mapping</i> .		
Total colocados seleccionados: 23	Frecuencia máxima: 18	Frecuencia mínima: 3
<ul style="list-style-type: none"> - Sustantivos, total 9: <i>address 18, technology 12, security 9, label 8, memory 5, hardware 4, name 4, function 3, software 3.</i> - Adjetivos, total 7: <i>efficient 5, explicit 4, good 4, dimensional 3, feasible 3, optimal 3, second 3.</i> - Siglas, total 6: <i>PCA 9, GNU 3.</i> - Prefijos, total 1: <i>RE 3.</i> 		
b) <i>Mapping</i> + colocado.		
Total colocados: 40	Frecuencia máxima: 44	Frecuencia mínima: 3
<ul style="list-style-type: none"> - Sustantivos, total 27: <i>problem 10, method 8, edges 8, step 7, process 6, decisions 5, table 4, strategy 4, services 3, functions 3/function 3, approach 2.</i> - Siglas, total 2: <i>diffeserv 2.</i> 		

3.5 Identificación de colocados contiguos especializados en las líneas de concordancia.

Some folks will notice that the Route Server address mapping function is basically the same problem that the Simulation Using VHDL-AMS is to build an empirical mapping function, $B = F(P)$, based on a large number into a stable target one [5]. Figure 6 The shape of the mapping function. Due to the transformation function be in either a map, graph, list or report. Mapping. GIS or mapping functionality can be embedded within existing known as commutative harmonic analysis) consists on mapping functions on to functions on a dual group ; in tion between business models and systems is similar to mapping functions on components. This is generally ref

Concordancias de mapping function(s).

- a) Colocado + mapping:
- | | |
|----------------------------|---------------------------|
| <i>dimensional mapping</i> | <i>feasible mapping</i> |
| <i>function mapping</i> | <i>hardware mapping</i> |
| <i>label mapping</i> | PCA mapping |
| <i>memory mapping</i> | <i>security mapping</i> |
| <i>name mapping</i> | <i>software mapping</i> |
| <i>optimal mapping</i> | <i>technology mapping</i> |
| <i>path mapping</i> | |
- b) mapping + colocado:
- | | |
|----------------------------|------------------------|
| <i>mapping decisions</i> | <i>mapping process</i> |
| <i>mapping edges</i> | <i>mapping step</i> |
| <i>mapping function(s)</i> | <i>mapping table</i> |
| <i>mapping method</i> | |

4. AGRUPACIONES LÉXICAS O CLUSTERS.

4.1 Distribución de clusters.

Combinaciones	Volumen	Frecuencia máxima
2-cluster	74	104
3-cluster	67	23
4-cluster	9	4
5-cluster	9	4

4.2 Inclusión de combinaciones especializadas en unidades más amplias.

<p style="text-align: center;">address mapping address mapping table PCA mapping the PCA mapping problem mapping method security mapping method technology mapping method</p>	<p style="text-align: center;">technology mapping based technology mapping technology mapping method mapping step explicit mapping step</p>
--	---

4.3 Colocados contiguos de las combinaciones especializadas.

PRE +	COMBINACIÓN	+ POST
<i>URL-IP, server, name-to-, multicast, MAC, IP-to-MAC, LANs, IP</i>	<i>address mapping</i>	<i>function, table</i>
<i>two, three</i>	<i>dimensional mapping</i>	
<i>energy-</i>	<i>efficient mapping</i>	
<i>no</i>	<i>explicit mapping</i>	<i>step</i>
<i>bt</i>	<i>feasible mapping</i>	<i>choices</i>
<i>total</i>	<i>function mapping</i>	<i>time</i>
<i>description to,</i>	<i>hardware mapping</i>	<i>is discarded</i>
	<i>label mapping</i>	<i>arrives, includes, messages, received</i>
<i>systems, setting up</i>	<i>memory mapping</i>	<i>hardware</i>
<i>oxmail, number-to-zone, host</i>	<i>name mapping</i>	
<i>uses</i>	<i>optimal mapping</i>	

<i>proposes, first, show mapping</i>	<i>path mapping PCA mapping</i>	<i>does, estimates, provides problem</i>
	<i>security mapping</i>	<i>entry(ies), method</i>
<i>previous</i>	<i>software mapping</i>	<i>alternatives, decision</i>
<i>based, does, study of, FPGA, conventional, BDD-based</i>	<i>technology mapping</i>	<i>approach, method</i>
<i>software, previous, represent, takes, due to</i>	<i>mapping decisions</i>	<i>imply, implicitly</i>
<i>activated, output, outgoing</i>	<i>mapping edges</i>	<i>specify</i>
<i>address, empirical, non-linear</i>	<i>mapping function(s)</i>	
<i>security, technology</i>	<i>mapping method</i>	
	<i>mapping process</i>	<i>includes</i>
<i>explicit, previous</i>	<i>mapping step</i>	<i>binds</i>
<i>address</i>	<i>mapping table</i>	<i>AMT, maintained</i>

8. Familia de *shell*. Forma analizada: SHELL

1. FRECUENCIA.

KEYWORD	Frec. Telec	Frec. Lacell	Ratio	Término, Chung	KEYNESS	P valor
908. SHELL	449	534	3,19334314	NO	302,5	0
Miembros relacionados:	F Telec	F Gene	Ratio	Término	Keyness	P valor
SUBSHELL	12	0	inf/esp	inf/esp	37,6	0
AEROSHELL	4	0	inf/esp	inf/esp	12,5	0,000397
Familia técnica: No, estadística y formalmente.						

2. DISTRIBUCIÓN.

Distribución por áreas	Distribución Keyness	Keykeyword	Nº de textos	Porcentaje
9	0	SHELL	15	0,91%
<i>Dispersion plot</i>				
Area	Words	Hits	Per 1,000	
2 Ar. Comp	329.643	83	0.25	
3 Telematics	1.205.064	297	0.25	
5 Materials	101.241	12	0.12	
7 Systems	307.691	7	0.02	
1 Electronics	722.823	14	0.02	
4 Signal proc.	580.936	11	0.02	
082 Esp.Tele	997.727	17	0.02	
6 Business	373.079	5	0.01	
081 Esp. Sign	867.208	4	0.00	

3. COLOCACIÓN.

3.1 Colocados.

N	Collocates	TOTAL	LEFT	RIGHT	L5	L4	L3	L2	L1	*	R1	R2	R3	R4	R5
1	THE	487	307	180	46	40	18	77	126	0	37	42	35	34	32
2	A	141	77	64	13	5	8	21	30	0	14	9	13	7	21
3	TO	113	54	59	15	13	15	11	0	0	8	10	11	18	12
4	OF	111	75	36	4	6	19	39	7	0	1	3	8	12	12
5	IN	101	60	41	15	13	17	9	6	0	12	8	5	9	7
6	AND	87	34	53	10	6	6	9	3	0	16	11	6	11	9
7	IS	70	26	44	13	3	8	2	0	0	10	11	4	10	9
8	FOR	48	27	21	4	4	5	6	8	0	4	2	4	5	6
9	YOU	44	17	27	2	9	5	0	1	0	7	6	4	6	4
10	YOUR	41	32	9	1	2	2	3	24	0	5	1	1	1	1
11	IT	37	13	24	2	5	5	1	0	0	4	4	4	8	4
12	THAT	37	18	19	8	6	2	1	1	0	4	3	9	2	1
13	SURFACE	34	6	28	1	2	2	1	0	0	5	5	6	6	6
14	OR	33	14	19	1	0	1	8	4	0	10	1	3	4	1
15	FROM	32	22	10	6	2	6	8	0	0	4	0	2	2	2
16	AN	30	10	20	2	1	4	3	0	0	9	3	4	1	3
17	THIS	29	13	16	6	0	4	2	1	0	4	2	2	5	3
18	UNIX	29	19	10	1	1	4	0	13	0	0	3	3	1	3
19	ARE	28	15	13	7	4	3	0	1	0	5	1	0	5	2

20	C	28	21	7	0	6	0	0	15	0	1	4	0	1	1
21	BE	26	8	18	1	4	2	1	0	0	6	4	4	1	3
22	BY	26	16	10	4	3	1	7	1	0	1	3	4	1	1
23	COMMAND	26	17	9	6	0	5	4	2	0	0	1	6	2	0
24	CAN	25	9	16	2	4	2	1	0	0	3	6	1	1	5
25	OPTIMIZED	25	22	3	4	5	4	4	5	0	2	1	0	0	0
26	WITH	25	10	15	4	2	1	3	0	0	2	2	2	7	2
27	AS	24	12	12	2	1	8	1	0	0	3	2	5	2	0
28	LOGIN	24	14	10	3	2	1	0	8	0	0	5	2	1	2
29	HOW	22	7	15	5	0	1	1	0	0	0	10	1	3	1
30	PROGRAM	22	9	13	1	2	2	4	0	0	0	2	3	3	5
31	USING	20	5	15	0	1	1	1	2	0	8	4	1	2	0
32	WILL	20	8	12	5	3	0	0	0	0	3	3	2	2	2
33	I	18	4	14	0	0	4	0	0	0	1	1	0	11	1
34	LINEAR	18	10	8	4	2	1	2	1	0	0	0	3	2	3
35	DO	17	2	15	1	1	0	0	0	0	3	2	9	0	1

3.2 Colocados significativos.

A) Colocados significativos en el corpus técnico.

Nº	Colocados	Z	T	MI
1	BOURNE	135,86	3,8698	10,267
2	TCSH	90,534	3,1584	9,6824
3	SETUID	54,276	2,4445	8,9454
4	USR	44,775	2,4422	8,3928
5	LOGIN	84,829	4,8827	8,2376
6	COMP	54,627	3,4503	7,9696
7	CYLINDRICAL	46,392	2,9876	7,9137
8	DELPHI	37,075	2,4389	7,8523
9	CSHRC	33,633	2,2263	7,8344
10	CSH	33,994	2,4369	7,6043
11	DEBUGGING	41,842	3,9641	6,7997
12	QUOTES	22,832	2,215	6,7313
13	BIN	24,938	2,4263	6,723
14	NETWARE	26	2,7957	6,4344
15	SCRIPTS	25,219	2,7937	6,3485
16	OPTIMIZED	41,886	4,9307	6,1731
17	CHILD	19,166	2,4107	5,9819
18	PROMPT	17,093	2,1991	5,9168
19	TC	17,007	2,1987	5,9027
20	VALENCE	15,071	2,1889	5,5669
21	COMMANDS	22,944	3,6471	5,3065
22	KERNEL	18,749	3,0771	5,2142
23	INTERPRETER	13,011	2,1737	5,163
24	THICKNESS	16,339	2,7484	5,1432
25	ANNUAL	12,591	2,1697	5,0735
26	HISTORY	16,849	2,9105	5,0666
27	EXECUTE	16,768	2,9097	5,0535
28	SHAPE	15,107	2,8902	4,7719

B) Colocados significativos en el corpus general.

Nº	Colocados	Z	T	MI
1	PDR	110,86	2,24	11,26
2	TORTOISE	82,41	2,83	9,73
3	FUELS	77,24	3,60	8,85
4	SNAIL	40,68	2,44	8,12
5	HOLLOW	33,91	2,81	7,19
6	TAKE	31,14	2,63	7,14
7	SHOCKED	31,18	2,97	6,78
8	SUITS	31,62	3,13	6,67
9	LIQUID	25,53	2,96	6,22
10	PRESERVED	20,48	2,60	5,95
11	CM	13,83	2,18	5,33
12	OUTER	13,77	2,18	5,32

3.3 Definición del nodo en el diccionario.

1. Significado de *shell* en la *General Service List* (West, 1953):

Shell (noun):

1. *Sea shell, snail shell, egg-shell, nutshell.*
2. *A 12-lb shell (war); shell-fire.*
3. *(various thin hard containers, casings, coverings) The shell of a boat, of a burned-out building, etc.*

Shell (verb): *Shell peas; shell the enemy.*

2. Definición de *shell* en el *Advanced Learner's Dictionary* de Cambridge (2003):

Shell (noun):

1. *(covering) The hard outer covering of something, especially nuts, eggs and some animals.*
2. *(explosive) A container, usually with a pointed end, which is filled with explosives and fired from a large gun.*
3. *(boat) A type of boat used for racing, powered by people using oars.*

Shell (verb):

1. *To remove peas, nuts, etc. from their shells or their natural protective covering.*
2. *To fire shells at something.*

3. Definición de *shell* en el diccionario especializado Webster:

Shell (noun),

Specialty Definition.

In computing:

1. *The command interpreter used to pass commands to an operating system; so called because it is the part of the operating system that interfaces with the outside world.*
2. *More generally, any interface program that mediates access to a special resource or server for convenience, efficiency, or security reasons; for this meaning, the usage is usually 'a shell around' whatever. This sort of program is also called a 'wrapper'.*
3. *A skeleton program, created by hand or by another program (like, say, a parser generator), which provides the necessary incantations to set up some task and the control flow to drive it (the term driver is sometimes used synonymously). The user is meant to fill in whatever code is needed to get real work done. This usage is common in the AI and Microsoft Windows worlds, and confuses Unix hackers.*

In electrical engineering:

A transparent or translucent gas-tight envelope enclosing the luminous element.

3.4 Colocados contiguos.

a) Colocado + shell.		
Total colocados seleccionados: 20	Frecuencia máxima: 13	Frecuencia mínima: 3
<ul style="list-style-type: none"> - Sustantivos, total 9: <i>Unix 13, Bourne 12, longin 8, elentron 7, netware 6, default 5, setuid 5, valence 5, Windows 3.</i> - Adjetivos, total 5: <i>cylindrical 8, secure 7, current 7, interactive 5, standard 3.</i> - Participios verbales, total 2: <i>debugging 5, optimized 5.</i> - Formas abreviadas, total 4: <i>C 15, comp 6, TC 3, TCSH 3.</i> 		
b) shell + colocado.		
Total colocados seleccionados: 13	Frecuencia máxima: 26	Frecuencia mínima: 3
<ul style="list-style-type: none"> - Sustantivos, total 12: <i>script 26/scripts 13, extensions 15/extension 12, optimization 13, variable 10/variables 3, command 7, program 7, prompt 3, shape 3, thickness 3.</i> - Prefijos, total 1: <i>mid 5.</i> 		

3.5 Identificación de colocados contiguos especializados en las líneas de concordancia.

sis [9]. In the first part of this paper, the linear elastic **shell optimization** process is presented in conjunction and B-spline surface representations may be used for **shell optimization** and the same surface can be expressed as a symmetric quadrant of the shell. Two cases of **shell optimization** are carried out using two types of *dramming* or *SQP* algorithm is employed in the present **shell optimization** system. In the second part of this paper two processes are involved: (a) the shape and sizing **shell optimization** and (b) the post-optimization, *geomatches*. In order to retain a prismatic shape during the **shell optimization** process, linking rules are required.

2.7.2. Linking of the design variables. During the **shell optimization** process, the linking of design variables is crucial in shell optimization. For **shell optimization**, various quantities have been used to interpolate any scalar quantity and for **shell optimization**, it can be used to interpolate the *x, y, z* coordinates. In this study, *SQP* is used.

2.6. Algorithm for **shell optimization**. The basic algorithm of the shape/sizing **shell optimization**. The selection of the design variables is crucial in **shell optimization**. For shell optimization, various quantities of the Coons patches serve as the design variables in **shell optimization**.

2.7.2. Linking of the design variables. Fig. 1 illustrates examples of various types of **shell optimizations**. However, a word of caution is ap

Concordancias de shell optimization(s).

a) Colocado + shell:

Bourne shell	netware shell	login shell
C shell	optimized shell	TCSH shell
cylindrical shell	secure shell	unix shell
debugging shell	setuid shell	valence shell
electron shell	standard shell	Windows shell
interactive shell	TC shell	

b) shell + colocado:

shell command(s)	shell prompt
shell extension(s)	shell script(s)
shell mid-	shell shape
shell optimization	shell variable(s)
shell program	

4. AGRUPACIONES LÉXICAS O CLUSTERS.

4.1 Distribución de clusters.

Combinaciones	Volumen	Frecuencia máxima
2-cluster	60	34
3-cluster	61	25
4-cluster	34	4
5-cluster	28	4

4.2 Inclusión de combinaciones especializadas en unidades más amplias.

debugging shell
debugging shell extensions
setuid shell
setuid shell scripts
shell optimization
shell optimization process
shell variable(s)
pass shell variable
shell variable settings

4.3 Colocados contiguos de las combinaciones especializadas.

PRE +	COMBINACIÓN	+ POST
<i>older, basic, use</i>	Bourne shell	<i>script(s)</i>
<i>use, unsupported</i>	<i>C shell</i>	<i>allows, creates, csh, uses, script</i>
	<i>cylindrical shell</i>	<i>example, modelled, roof, subjected</i>
	debugging shell	<i>extensions</i>
<i>first</i>	<i>electron shell</i>	
<i>allow, non-</i>	<i>interactive shell</i>	<i>escapes, prompt</i>
<i>changing, username</i>	login shell	
<i>including</i>	netware shell	<i>packages, runs, supported</i>
<i>non-, thickness-, shape-</i>	optimized shell	<i>obtained, illustrated</i>
<i>SSH, original, authored, Solaris</i>	<i>secure shell</i>	<i>implementation, protocols, software, SSH</i>
<i>get</i>	setuid shell	scripts
	TC shell	<i>called, tcsh</i>
<i>each</i>	TCSH shell	<i>alias, redirect</i>
	<i>unix shell</i>	<i>maintains, provides, scripts</i>
<i>outer, full, own</i>	valence shell	
<i>developing, start</i>	<i>Windows shell</i>	<i>automatically, extensions</i>
<i>each, built-in, system, basic, entering</i>	<i>shell command(s)</i>	<i>instructs, consists</i>
<i>every, registered, debugging, Windows, write, successful</i>	<i>shell extension(s)</i>	debugging , <i>code, work, developing</i>
	<i>shell mid-</i>	<i>surface</i>
<i>elastic, present, sizing</i>	<i>shell optimization</i>	<i>process, system</i>
<i>remote, same</i>	<i>shell program</i>	<i>rsh, starts</i>
<i>Non-interactive</i>	<i>shell prompt</i>	<i>named, usr, does,</i>
<i>C-, little, Bourne, setuid, Unix, executable, writing</i>	<i>shell script(s)</i>	
	<i>shell shape</i>	<i>optimization</i>
<i>history, pass, using</i>	<i>shell variable(s)</i>	<i>called, settings, cwd, old, depends</i>

9. Familia de *hub*. Forma analizada: HUB

1. FRECUENCIA.

KEYWORD	Frec. Telec	Frec. Lacell	Ratio	Término, Chung	KEYNESS	P valor
699. HUB	245	79	11,7782155	NO	445,4	0
Miembros relacionados:	F Telec	F Gene	Ratio	Término	Keyness	P valor
HUBS	97	34	10,8351099	NO	170,1	0
SMARTHUB	5	0	inf/esp	inf/esp	15,7	0,000075
Familia técnica: No, estadística y formalmente.						

2. DISTRIBUCIÓN.

Distribución por áreas	Distribución Keyness	Keykeyword	Nº de textos	Porcentaje
7 (- 2, 5)	802	HUB	24	1,45%
<i>Dispersion plot</i>				
Area	Words	Hits	Per 1,000	
082 Esp.Tele	997.727	130	0.13	
3 Telematics	1.205.064	58	0.05	
081 Esp. Sign	867.208	35	0.04	
6 Business	373.079	8	0.02	
4 Signal proc.	580.936	8	0.01	
1 Electronics	722.823	6	0.01	
7 Systems	307.691	1	0.00	

3. COLOCACIÓN.

3.1 Colocados.

N	Collocates	TOTAL	LEFT	RIGHT	L5	L4	L3	L2	L1	*	R1	R2	R3	R4	R5
1	THE	188	124	64	16	19	12	17	60	0	0	27	15	8	14
2	A	151	111	40	10	14	15	17	55	0	0	17	8	3	12
3	TO	78	46	32	11	8	10	15	2	0	6	2	8	13	3
4	AND	70	27	43	5	3	7	6	6	0	22	9	4	2	6
5	OF	52	24	28	3	4	3	11	3	0	8	0	2	5	13
6	OR	50	29	21	4	4	2	5	14	0	8	7	3	1	2
7	IS	40	13	27	5	5	1	2	0	0	6	3	7	4	7
8	AS	29	17	12	1	3	3	10	0	0	2	2	2	3	3
9	IN	29	16	13	5	1	4	6	0	0	4	0	2	3	4
10	NETWORK	28	14	14	2	2	3	4	3	0	2	3	2	6	1
11	AN	25	12	13	0	0	5	7	0	0	0	6	3	2	2
12	ARE	23	7	16	1	5	0	1	0	0	0	2	5	3	6
13	SWITCH	23	9	14	1	1	1	6	0	0	1	5	5	1	2
14	WITH	22	7	15	0	1	2	3	1	0	8	4	1	1	1
15	FROM	19	15	4	2	2	4	7	0	0	1	1	1	0	1
16	THAT	17	5	12	1	1	1	2	0	0	3	4	1	1	3
17	BE	16	8	8	3	4	1	0	0	0	0	3	1	3	1
18	EACH	16	8	8	2	2	1	0	3	0	0	4	2	0	2
19	SPOKE	16	4	12	1	0	2	1	0	0	0	11	0	1	0
20	FOR	15	4	11	2	2	0	0	0	0	3	3	0	3	2
21	ON	15	5	10	0	1	1	3	0	0	3	1	1	4	1

22	AT	14	14	0	4	0	2	8	0	0	0	0	0	0	0
23	CENTRAL	14	9	5	1	0	0	2	6	0	0	0	2	0	3
24	ROUTER	14	4	10	1	2	0	0	1	0	3	1	2	2	2
25	HAS	12	3	9	1	0	1	1	0	0	4	3	1	0	1
26	INTO	12	9	3	0	5	0	3	1	0	0	1	1	1	0
27	NODE	12	3	9	0	0	2	1	0	0	8	1	0	0	0
28	BASED	11	2	9	1	0	1	0	0	0	9	0	0	0	0
29	CONNECTED	11	6	5	2	1	3	0	0	0	0	1	3	1	0
30	PORT	11	6	5	0	0	1	0	5	0	1	0	2	2	0
31	ALL	10	1	9	0	0	0	0	1	0	1	2	1	3	2
32	I	10	4	6	3	0	1	0	0	0	0	4	1	0	1
33	IT	10	3	7	0	1	0	0	2	0	1	2	1	0	3
34	LAN	10	7	3	1	0	4	0	2	0	1	0	1	0	1
35	STATION	10	7	3	4	2	0	0	1	0	0	0	1	1	1

3.2 Colocados significativos.

A) Colocados significativos en el corpus técnico.

Nº	Colocados	Z	T	MI
1	STAR	17,95	2,20	6,05
2	SPOKE	111,97	3,99	9,62

* Otros colocados significativos, pero que no cumplen todos los requisitos:

Nº	Colocados	Z	T	MI	Nº	Colocados	Z	T	MI
1	HEAD	17,68	2,59	5,54	16	COMPUTERS	6,55	2,02	3,39
2	CENTRAL	23,69	3,65	5,39	17	NODE	9,62	3,10	3,26
3	MESH	14,03	2,18	5,37	18	CABLE	7,16	2,36	3,20
4	DIRECTED	12,99	2,17	5,16	19	LOCATION	5,80	1,98	3,10
5	INTERMEDIATE	13,14	2,37	4,94	20	SWITCHING	5,79	1,98	3,10
6	COMMENT	11,76	2,16	4,89	21	ETHERNET	6,72	2,33	3,06
7	ATTACHED	13,26	2,55	4,76	22	ROUTER	9,33	3,28	3,01
8	SWITCH	23,05	4,60	4,65	23	EQUIPMENT	4,48	1,85	2,55
9	MODEM	13,59	2,72	4,65	24	COMMON	4,29	1,83	2,46
10	STATION	13,65	3,01	4,36	25	POINT	5,21	2,37	2,26
11	PORT	13,23	3,13	4,16	26	NODES	3,87	1,77	2,26
12	CONNECTED	12,72	3,12	4,06	27	END	4,80	2,22	2,22
13	LAN	11,54	2,95	3,93	28	DEVICES	4,45	2,07	2,21
14	TOPOLOGY	7,75	2,08	3,80	29	CONNECTION	3,74	1,75	2,20
15	SITE	8,64	2,58	3,49	30	COMMUNICATIONS	3,91	1,88	2,11

B) Colocados significativos en el corpus general.

Nº	Colocados	Z	T	MI
1	ETHERNET	134,033516	2,23544598	11,81177207
2	MICRO	92,8225504	2,44778634	10,48984398
3	PORT	49,156631	2,81912411	8,248086233
4	CARD	25,0996908	2,42659352	6,741228025

3.3 Definición del nodo en el diccionario.

1. Definición de *hub* en el *Advanced Learner's Dictionary* de Cambridge (2003):

Hub (noun):

1. *The central or main part of something where there is most activity: The City of London is the hub of Britain's financial world.*

2. *The central part of a wheel into which the spokes (= bars connecting the central part to the outer edge of the wheel) are fixed.*

2. Definición de *hub* en el diccionario especializado Webster:

Hub (noun):

1. *The central part of a car wheel (or fan or propeller etc) through which the shaft or axle passes.*
2. *The center of a city.*

Specialty Definition:

In computing:

A device connected to several other devices. In ARCnet, a hub is used to connect several computers together. In a message handling service, a number of local computers might exchange messages solely with a hub computer. The hub would be responsible for exchanging messages with other hubs and non-local computers.

In electrical engineering: A circular device on which the magnetic tape is wound.

3.4 Colocados contiguos.

a) Colocado + <i>hub</i> .		
Total colocados seleccionados: 7	Frecuencia máxima: 6	Frecuencia mínima: 2
<ul style="list-style-type: none"> - Sustantivos, total 4: <i>port 5, communications 4, network 2, Ethernet 2.</i> - Adjetivos, total 1: <i>central 6.</i> - Participios verbales, total 1: <i>switching 2.</i> - Formas abreviadas, total 1: <i>LAN 2.</i> 		
b) <i>Hub</i> + colocado.		
Total colocados seleccionados: 7	Frecuencia máxima: 8	Frecuencia mínima: 2
<ul style="list-style-type: none"> - Sustantivos, total 5: <i>node 8, location 5, site 4, router 3, equipment 2.</i> - Participios verbales, total 1: <i>based 9.</i> - Formas personales de verbos, total 1: <i>see 2.</i> 		

3.5 Identificación de colocados contiguos especializados en las líneas de concordancia.

*ance of 40 km between nodes at 622 Mb/s, with a **hub node** and two access nodes. The ring employ bed consists of a unidirectional ring with a central **hub node** and many access nodes. It uses six wa d with fixed wavelength conversion capability at its **hub node** that can support all lightpath sets with l r a star network with nodes, even, and where the **hub node** has FCWP. Then any request with load ns. Each unidirectional link has channels at . The **hub node** has no conversion. Lightpaths are unidi channel if its unidirectional channel going into the **hub node** has wavelength (). Thus, bidirectional could have for and . A star network consists of a **hub node**, one or more rim nodes, and all links ar d to a wavelength in (if it is directed into (out of) a **hub node**. Such an assignment is possible becau ched at its hub when $W = 4$. (a) Star network. (b) **Hub node**. Fig. 13. (a) Request $fp_0; p_1$; for a sta are incident to the hub (see Fig. 12). We consider **hub nodes** with fixed wavelength conversion such*

Concordancias de *hub node(s)*.

a) Colocado + *hub*:

central hub
port hub
communications hub
network ub
Ethernet hub
LAN hub
switching hub

b) *Hub* + colocado:

hub based

hub equipment
hub location
hub network
hub node
hub router
hub site

4. AGRUPACIONES LÉXICAS O CLUSTERS.

4.1 Distribución de clusters.

Combinaciones	Volumen	Frecuencia máxima
2-cluster	20	8
3-cluster	33	11
4-cluster	9	3

4.2 Inclusión de combinaciones especializadas en unidades más amplias.

<p><i>port hub</i> <i>a (nº) port hub</i> communications hub <i>intermediate communications hub</i></p>

4.3 Colocados contiguos de las combinaciones especializadas.

PRE +	COMBINACIÓN	+ POST
<i>common</i>	<i>central hub</i>	<i>called, node</i>
<i>intermediate</i>	<i>communications hub</i>	
<i>existing</i>	<i>Ethernet hub</i>	
<i>AT&T's</i>	<i>network hub</i>	
<i>uplink, four-</i>	<i>port hub</i>	<i>linked, router</i>
<i>intelligent, auto</i>	<i>switching hub</i>	
<i>failure-, making</i>	<i>hub- based</i>	<i>network(s), LANs</i>
<i>rigid</i>	<i>hub-and-spoke</i>	<i>topology, design,</i> <i>networks, internetwork</i>
<i>central</i>	<i>hub node</i>	
<i>four-port, dedicated</i>	<i>hub router</i>	
<i>LMDS, alternate</i>	<i>hub site(s)</i>	<i>equipment</i>

10.6 Conclusión.

Las muestras del lenguaje recogidas en el corpus especializado se han ido sometiendo a una serie de pruebas, que han hecho posible clasificar las unidades lingüísticas de forma semiautomática en tres categorías: especializadas, académicas y generales.

El procesamiento del corpus ha permitido efectuar una clasificación del vocabulario según criterios estadísticos y formales suficientemente satisfactoria, facilitando en gran medida el acceso a un conglomerado de datos y proporcionando, a la vez, las pistas necesarias para encauzar un análisis cualitativo.

Los datos obtenidos del análisis detallado constituyen una fuente de información fundamental para la determinación de la naturaleza de las formas examinadas y la caracterización del perfil léxico de las telecomunicaciones. Esta información revela el comportamiento de las unidades léxicas en el contexto donde se activan sus significados, de acuerdo con los rasgos pragmáticos que definen el sublenguaje especializado al que pertenecen.

El estudio cualitativo de las formas pertenecientes a las diferentes categorías, fundamentado y combinado con información cuantitativa, ha llevado a las conclusiones que se exponen a continuación.

El primer conjunto de formas analizadas (apartado 10.3) engloba familias de palabras en las que todos o más de la mitad de los miembros se valoran como especializados, de tal modo que se consideran familias técnicas. El análisis detallado ha corroborado el carácter especializado determinado estadísticamente, a la vez que ha revelado el comportamiento léxico característico de los representantes de cada familia.

Todas las formas analizadas se combinan con otras unidades lingüísticas dando lugar a construcciones que reflejan nociones y conceptos propios del dominio. Incluso hay formas que llegan a constituir clusters especializados de hasta cinco componentes (*network layer **protocol** configuration negotiation, smartpartner pager **wireless** data service, reliable multicast over **satellite networks**, next generation **satellite** system NGSS, transmission line characteristic **impedance** values*). Asimismo, es frecuente encontrar colocados significativos como integrantes de construcciones técnicas, que muestran patrones de comportamiento léxico típico y en numerosas ocasiones distintivos del dominio, ya que la forma bajo análisis no aparece en el lenguaje general (*optical crosstalk, Spice netlist, analyser applet, shielded suspended microstrip, tree-based reliable multicast*).

La identificación de los colocados significativos ha resultado ser un elemento fundamental, puesto que contribuye a desvelar el entorno semántico típico con el que se relaciona una forma y el sentido con el que ésta se ha utilizado. De hecho, los colocados significativos de cada forma están relacionados con la definición de ésta en el diccionario técnico, ayudan a transmitir el significado especializado y reflejan su uso específico. *Netlist*, *Crosstalk* y *timeslot* suponen una excepción, en el sentido de que no son suficientemente frecuentes para poder hallar sus colocados significativos, y presentan un número menor de combinaciones especializadas. Sin embargo, su carácter específico queda enfatizado por el hecho de que tales formas no aparecen en el corpus general, no se distribuyen por todas las áreas de conocimiento, el resto de miembros de la familia también son especializados y además, no están registradas en el diccionario general pero sí en el técnico.

Todas las formas técnicas analizadas comparten tres rasgos. Por un lado, han alcanzado el estatus de *key-keyword*, en mayor o menor grado, es decir, su incidencia es significativa en comparación con el lenguaje general y entre las diferentes secciones del corpus de telecomunicaciones, además de ser palabra clave de una a tres áreas de conocimiento (salvo *timeslot*). Por otro lado, todas ellas se combinan con formas truncadas (excepto *latency*, *impedance* y *applet*) dando lugar a construcciones especializadas y siendo usual que integren la propia sigla (*CDP: Cisco discovery protocol; WLAN: wireless local area network; MVR: multiple chip rate; DCT: discrete cosine transform; UTD: uniform theory of diffraction*). Las formas abreviadas evidencian el carácter especializado del registro y el conocimiento específico requerido para poder entender este sublenguaje. Asimismo, las familias de un solo miembro (*Bluetooth*, *cosine*, *Crosstalk*) han demostrado ser altamente especializadas, al igual que aquellas cuyo miembro representante no está catalogado en el diccionario general (*Bluetooth*, *multicast*, *netlist*, *impedance*, *Crosstalk*, *microstrip*, *timeslot*). Con frecuencia, las acepciones del diccionario general muestran un significado totalmente distinto al registrado en el diccionario técnico, el cual suele contemplar el uso de la palabra en uno o varios subdominios de las telecomunicaciones.

Cabe destacar el comportamiento de *chip* y de *satellite*, dado que son formas valoradas estadísticamente como no-términos, al mismo tiempo que son representantes de una familia clasificada en conjunto como especializada, ya que más de la mitad de los miembros son evaluados como específicos. El análisis detallado demuestra que *chip* y *satellite* son términos por los colocados significativos que atraen y definen su entorno semántico. Además, por lo que respecta a *chip*, las acepciones registradas en el diccionario general y en el técnico son totalmente distintas, coincidiendo el significado y uso de *chip* en el diccionario técnico con los colocados significativos hallados en el corpus. Finalmente, *chip* y *satellite* se combinan con otras unidades léxicas (*flip chip*,

mobile satellite, chip assembly, satellite operator), con formas abreviadas (*LAN chip, NGSO satellite communications, FPGA chip, VSAT satellite technology*) e integra formas truncandas (*VCR: variable chip rate, DBS: direct broadcast satellite, SoC: system-on-a-chip, satellite terminals*).

El segundo grupo de palabras analizadas (apartado 10.4) incluye aquellas formas que, aun perteneciendo a la lista de vocabulario académico, podían haber sufrido un cambio o adaptación del significado en el registro específico.

Existen diversos factores que conducen al replanteamiento de la condición académica de *network, data, router* y *layer*. En primer lugar, se encuentran entre las 15 familias de palabras más significativas de las telecomunicaciones (*network* es la primera y *data* la segunda). Esto quiere decir que, en comparación con un corpus representativo de la lengua general provisto de textos académicos, la presencia de tales formas es significativa de forma sobresaliente, cuando la frecuencia de uso debería ser aproximada. Además, estas formas están relacionadas cualitativamente con las telecomunicaciones, constituyen conceptos fundamentales e incluso llegan a denominar ramas del dominio principal (*communication networks and systems; communication networks services; data transmission; computer networks*).

En cuanto al resto de miembros de la familia que representan, más de la mitad no se contemplan en el listado académico de Coxhead, a la vez que son calificados como términos de acuerdo con los criterios de Chung.

El análisis detallado revela cómo las palabras académicas analizadas se combinan e integran numerosas formas abreviadas (*MWTN: multiwavelength transport network; LSR: label switch router; SDF: synchronous data flow; VR: virtual router; CDPD: cellular digital packet data*). Algunas siglas han experimentado un proceso de nominalización, derivando en formas sustantivadas de uso extendido (*LAN: local area network; ADSUs: application data units; WAN: wireless area network; PLP: presentation layer protocol; SSL: secure socket layer*).

Estas formas académicas presentan un alto rendimiento, originando un elevado número de combinaciones que forman unidades léxicas mayores y clusters de diferentes tamaños y de carácter especializado. Entre las construcciones resultantes se encuentran aquellas compuestas por unidades léxicas del lenguaje general que adquieren significados nuevos mediante el proceso de analogía, y cuyo significado es deducible del significado de las partes. Las características de tales construcciones llevan a clasificarlas como vocabulario semitécnico (*star network, global network, data access, voice data integration, conventional router, boundary layer, link layer, transport layer security*). Por otro lado, también se encuentran construcciones altamente especializadas, puesto que se precisa conocimiento sobre la materia para conseguir una comprensión

plena (*ATM network, AppleTalk network, switched multimegabit data service SMDS, OSPF data structure, multiplexed data stream, configurable router interface, label switch router, OSI presentation layer, network layer protocol configuration, Ethernet physical layer*).

Todas las particularidades señaladas sobre el comportamiento de las formas académicas llevan a reconsiderar su condición y a replantear la decisión adoptada inicialmente de asumir lo siguiente: las palabras clave identificadas en el corpus no estarán provistas de significados adicionales, siempre que coincidan con las formas registradas en el listado académico de Coxhead (2000). Sin embargo, la información obtenida sobre el comportamiento de tales formas en el corpus técnico requiere añadir una condición más: las palabras clave deben coincidir con las formas registradas en la lista académica y además deben utilizarse individualmente. Esto quiere decir que la forma mantiene su carácter académico cuando no se combina con otras unidades léxicas, de lo contrario, constituye construcciones especializadas, transformando su condición académica en especializada. Las siguientes líneas de concordancia ilustran el uso académico de las formas:

<p>... agent traverse the network from the source to destination... ... doing so makes the network unnecessarily complex... ... a message destined for your network by formatting... ... What's the hardest part of the data analysis? further observe that data extracted in earlier passes can... ... satellites allow scientists to gather valuable data about the ...</p>

El último grupo de formas analizadas (apartado 10.5) recoge los representantes de las familias clasificadas estadística y formalmente como generales, pero que cualitativamente y a juicio de la analista, podrían haber experimentado un cambio o adaptación del significado al área especializada.

La observación y el estudio del comportamiento de las muestras analizadas confirman las dudas sobre la condición de las palabras seleccionadas. A excepción de *banking* y *tunning*, el resto de formas muestran un perfil específico distintivo en el corpus técnico. Entre ellas se distinguen dos tipos de comportamiento: formas que adquieren un significado especializado por sí mismas (*noise, linear, internet, programming, host, bus, mapping, shell, Windows*) y formas que adquieren un significado especializado al combinarse con otras unidades léxicas, dando lugar a construcciones especializadas (*signal, memory, scenario, hub*).

En conjunto, comparten las siguientes características:

- Se utilizan en el corpus técnico con un sentido diferente al registrado en el listado de West, exceptuando a *hub* que no se halla en dicha lista.

- Se distribuyen por las nueve áreas de conocimiento que integran la ingeniería de telecomunicación, a excepción de *scenario* y *hub* que se distribuyen por ocho y siete áreas respectivamente.
- Aun siendo palabras clave en el conjunto del corpus, no se manifiestan como palabra clave de un área en concreto, salvo *linear*, *hub* y *mapping*.
- Constituyen construcciones especializadas, esenciales para expresar adecuadamente el conocimiento especializado.
- Integran y se combinan con formas truncadas, excepto *scenario* y *hub*.

Una vez más, los colocados significativos del nodo han demostrado ser un factor elemental para determinar la condición de las formas. Éstos evidencian el uso diferenciado de la unidad léxica en los dos registros y revelan si tal unidad ha adquirido un carácter especializado en el dominio.

Los criterios formales y estadísticos sobre los que se fundamenta la clasificación inicial, revelan un comportamiento similar de las formas generales frecuentes en los dos corpus. Sin embargo, el análisis detallado prueba que unidades léxicas del lenguaje general han sufrido una modificación o extensión del significado, convirtiéndose en términos propios de las telecomunicaciones. En ocasiones, los colocados significativos reflejan el doble uso de la forma en los corpus, es decir, que una misma forma se ha utilizado con el sentido general y con el específico (*memory*, *signal*, *windows*). Asimismo, debido a la globalización y difusión del conocimiento, el diccionario de la lengua general recoge la acepción especializada (*memory*, *noise*, *internet*, *linear*, *radio*, *programming*, *host*, *hub*).

La forma *windows* es un caso destacable. Los colocados significativos en el corpus técnico, las líneas de concordancia, las combinaciones especializadas y los clusters revelan el uso especializado asociado a *Windows*: el sistema operativo (escrito siempre con mayúscula y en plural). Por otro lado, en el corpus general se observan colocados significativos de *windows*, como plural de *window* (*clerestory*, *shutters*, *glazed*, *blinds*, *walls*, *curtains*, etc...) y colocados significativos de *Windows* (*Macintosh*, *wordperfect*, *IBM*, *versions*, *programs*, *interfaces*, etc...). En realidad, el programa ha computado dos formas diferentes como una (no diferencia mayúsculas y minúsculas en el listado de frecuencia), de modo que los valores obtenidos se refieren a la suma de las ocurrencias de una forma con dos sentidos distintos.

En definitiva, todas estas formas cuyo significado varía al integrarse en el dominio especializado, más aquellas que designan un concepto de forma unívoca, junto a las combinaciones que ambos tipos generan, constituyen el vocabulario especializado de la ingeniería de telecomunicación y definen su perfil léxico característico. Por lo tanto,

después de los análisis realizados, es posible concluir que las formas situadas entre las 1000 familias de palabras más significativas del dominio, valoradas como especializadas (Apéndice IX), deberían encontrarse en los libros de texto destinados a la enseñanza del inglés profesional y académico de la especialidad.

A continuación se anotan las 1.000 formas más significativas del vocabulario especializado de la ingeniería de telecomunicaciones (K: valor de *keyness*). Las 2.747 que constituyen el repertorio completo se encuentran en el Apéndice IX.

Forma especializada	K	Forma especializada	K	Forma especializada	K	Forma especializada	K
NETWORK	41785	SYNCHRONOUS	1017	LIGHTPATHS	430	DEBUG	273
DATA	31852	HOP	1017	MAPPINGS	429	SYNCHRONIZE	272
SIGNAL	17923	PROGRAMMABLE	1017	PLANAR	429	OCTET	272
SYSTEMS	17378	DVB	1013	DOPPLER	428	NETWARE	272
IP	16182	WAVEFORM	1007	CACHING	425	PROXIES	271
SYSTEM	14832	DYNAMICALLY	1007	ESTIMATOR	424	PID	271
PROTOCOL	13678	BEAM	1007	WLANS	420	EXTENSIBLE	270
ROUTER	11974	COMPUTE	999	PNS	420	SPATIALLY	270
WIRELESS	11454	MICROSTRIP	994	BACKPLANE	420	LOGIN	270
FREQUENCY	11440	XML	992	SUBNETWORK	417	BUS	270
ROUTING	10691	WINDOWS	992	SOLARIS	416	UDRP	270
LAYER	10605	SAMPLING	989	ROBUSTNESS	414	BUS	270
MODEL	9860	RADIO	952	WCDMA	414	MV	269
INTERFACE	9558	SPS	949	DATAFLOW	414	AAL	269
BANDWIDTH	9551	DNS	941	HRM	414	CE	268
PACKET	9485	THRESHOLD	935	TDM	414	UNREACHABLE	267
CIRCUIT	9349	AMPLIFIERS	930	SIMULINK	414	INTEROPERABLE	267
SERVER	8963	BOOLEAN	921	CABLING	414	PDA	266
SOFTWARE	8471	SONET	910	TELEPHONE	413	ION	266
SIMULATION	8189	FPGAS	900	STOCHASTIC	413	ISI	266
VOLTAGE	7743	SCHEMA	896	BUFFERING	411	SCL	266
OPTICAL	7658	PHONE	896	CRC	411	MEMS	266
TRAFFIC	7421	ENCODING	893	INFRASTRUCTURES	410	CW	264
ALGORITHM	7264	ANALOGUE	892	KB	409	KEYING	264
NODE	7042	FLows	890	SNR	408	MICROPAYMENTS	264
FILTER	6671	WAVEGUIDE	887	XP	405	VLR	264
PACKETS	6100	FIBERS	882	CONSTRUCTOR	405	RECONFIGURATION	263
NODES	5952	POINTER	877	RTT	405	MAXIMAL	263
ROUTERS	5876	THREADS	865	QPSK	402	AUTHENTICATED	263
WEB	5838	VECTORS	864	BANKING	398	DOPING	262
PROTOCOLS	5800	ROAMING	864	TRANSCEIVER	398	CONDUCTIVE	259
CIRCUITS	5718	ACK	860	LMS	398	CENTERS	258
SERVICE	5707	UDP	856	INTRANET	397	WANS	256
FIBER	5659	STREAM	852	OC	396	OPTIMIZING	256
ETHERNET	5602	SAMPLES	851	SPECTRA	394	FIBRE	255
SERVICES	5524	INTERCONNECT	850	WAVEFORMS	393	CAPACITIVE	253
TCP	5248	VLSI	847	TELNET	393	PARAMETERIZED	251
CONFIGURATION	5134	MATLAB	841	XPOWER	392	VOLTS	251
HARDWARE	4940	TM	830	ISPS	392	ALGORITHMIC	249
ALGORITHMS	4829	AC	828	ASIC	392	TOOLBAR	249
ATM	4817	MULTIPOINT	825	ADJACENCY	391	PROTOTYPING	249
TYPE	4717	PCS	822	AMP	391	NMOS	248
LAN	4387	MODEMS	818	ISO	390	EXTRANET	248

ARCHITECTURE	4326	TDMA	812	CONVOLUTION	390	THREADED	247
PATH	4197	MOS	808	GNU	390	CLUSTERS	247
OSPF	4027	CORBA	796	ENCAPSULATION	389	MODULAR	246
WAVELENGTH	4010	IC	795	TOOLKIT	388	HDLC	245
EXAMPLE	3993	KBPS	793	TOPOLOGICAL	387	PHOTODETECTOR	245
JAVA	3913	AGGREGATION	784	STATION	387	SOLVE	242
NUMBER	3800	SCALABLE	784	LOOKUP	386	FIRMWARE	241
SWITCH	3731	RECURSIVE	783	MIMO	386	SYNCHRONISATION	241
ANALOG	3715	THREAD	769	CONFORMANCE	386	DIRECTORIES	240
QOS	3623	SMS	765	CDR	384	PHOTODETECTORS	238
VHDL	3607	UTILIZATION	764	LISP	383	TEXTIO	238
ANTENNA	3598	CONCURRENCY	753	STREAMING	382	MULTIPLEXERS	235
LINEAR	3552	IETF	740	CONDUCTION	382	PIPELINING	235
MPLS	3488	SCENARIO	736	MULTIPLEXED	381	APPLETS	234
MODELS	3480	APPLET	734	CHECKSUM	381	PVC	233
SOLUTIONS	3476	INVERTING	727	SIZES	381	LOGGING	233
GSM	3427	RSVP	726	PARALLELISM	380	OPTOELECTRONIC	233
SPECTRUM	3418	RIP	716	CVD	380	INTERNETWORKING	233
LINUX	3377	PMI	715	OUTAGE	379	ADAPTERS	233
INTERFACES	3364	DECODER	708	HDL	378	EOMPLS	232
POWER	3324	OSCILLATOR	703	PIXEL	377	EXPONENTIALLY	230
PROGRAMMING	3309	MICROPAYMENT	703	CWND	376	MULTIPROTOCOL	229
REMOTE	3277	HANDOVER	696	NTP	376	TFRC	229
CELL	3253	DIODES	695	TLS	376	VIRTUALIZATION	228
CABLE	3245	SEMICONDUCTORS	694	SUBNETS	374	DECOUPLING	227
AUTHENTICATION	3234	NS	694	EDA	371	COMPARATOR	227
SERVERS	3232	POLARIZATION	692	SMART	368	STEPS	226
VPN	3098	INTERCONNECTION	690	AGILENT	367	PMOS	226
FILTERS	3090	LSR	687	SCATTERING	365	COVARIANCE	225
SWITCHING	3083	PN	686	TRANSCONDUCTANCE	364	CENTRIC	225
IEEE	3044	KHZ	685	GRANULARITY	364	WORKFLOW	225
BROADBAND	3010	MODULATED	684	SCWP	364	PHOTON	225
SOLUTION	2972	NSSA	681	SDH	364	MAILING	225
TOOLS	2968	OXIDE	679	DATAGRAMS	364	MULTICHANNEL	224
PARALLEL	2915	PPP	679	SEMISTRUCTURED	364	OPTIMISED	223
TYPES	2906	COSINE	679	ADJACENCIES	364	CACHED	223
NOISE	2887	MODELED	678	CHIPS	364	RETRANSMITTED	222
MEMORY	2743	RETRANSMISSION	678	INTERCONNECTS	363	TABLES	220
SATELLITE	2715	DISTANCE	671	HANDOFF	362	MULTICASTING	220
LSAS	2676	BER	670	LEDS	361	TESTABILITY	220
MODULATION	2672	EMAIL	669	ISM	361	TRANSCIVERS	217
SWITCHES	2646	APPLETALK	668	NBMA	361	BUFFERED	215
MULTICAST	2625	IOS	665	AMPS	360	INFINIBAND	213
DSP	2523	HTML	663	SN	360	DBM	213
LSA	2522	SCALING	662	WIRES	360	ACTIVATABLE	213
CDMA	2510	ICS	662	SIGE	357	WAVES	211
CISCO	2498	ADSL	661	SAMPLED	355	EMBEDDING	210
MODELING	2496	TOPOLOGIES	659	PROMELA	354	FAILOVER	210
IMPLEMENTATIONS	2463	SPATIAL	658	LUCENT	354	SOLVING	207
RELAY	2460	NUMERICAL	656	CODEC	353	GLOBALS	207
CELLULAR	2420	FCC	655	MULTIMODE	353	INTERFERERS	207
DIRECTORY	2411	PACKAGES	653	POINTERS	353	PBXS	207
TOPOLOGY	2408	REFRACTIVE	652	GAAS	352	NPID	207
CARRIER	2385	RUNTIME	650	RTP	351	HOSTING	207

SEMICONDUCTOR	2385	NMS	649	SUBLAYER	351	VSAT	206
GLOBAL	2367	HASHES	646	INDUCTOR	351	INDEPENDENT	206
LATENCY	2360	TTL	644	NOC	350	ELECTRODE	206
THROUGHPUT	2349	UMTS	640	QUANTIZATION	350	DIPOLES	205
CHIP	2340	ENCODER	637	CNC	349	SCHEDULERS	205
VENDOR	2335	TI	636	COMPUTES	349	MACROPAYMENT	204
AMPLIFIER	2322	FEC	634	OCTETS	348	LOGARITHMIC	203
MHZ	2320	FIR	631	AUTHENTICATE	348	PHOTODIODE	203
GHZ	2276	NT	627	RESONATOR	347	SUPERCONDUCTORS	203
CACHE	2272	EXAMPLES	626	DEPENDENCIES	345	INTERNATIONALIZING	200
FUZZY	2270	TMN	625	COMINT	345	CODECS	200
VECTOR	2238	VR	623	BIDIRECTIONAL	345	INITIALIZED	199
FPGA	2236	WIRED	622	LAPTOP	344	BANDS	199
FUNCTIONALITY	2210	GENERATOR	620	ALIASING	343	ENUM	198
SCTP	2205	CIRCUITRY	619	MIPS	343	PARALLELIZATION	198
FREQUENCIES	2188	SNA	618	PROFILING	343	STATICALLY	198
ELECTRONIC	2182	POINT	614	DWT	342	PIXELS	198
RF	2162	CENTER	608	PHASOR	342	WORKGROUP	197
DB	2150	UPSTREAM	607	DAC	341	OPTIONALLY	196
TYPICALLY	2134	CRYPTOGRAPHIC	606	FDR	340	INITIALIZE	196
WLAN	2123	DIALOG	606	CACHES	340	GBIT	195
ARCHITECTURES	2119	MAIL	604	CRYPTOGRAPHY	340	CDRS	195
CLIENT	2111	GENERATING	603	MULTIPLEXER	339	CMTS	195
INTERFERENCE	2093	DMCS	602	VRF	339	PHOTORESIST	194
MODULE	2090	VLAN	602	TABLET	338	LOWPASS	194
ISDN	2061	COMPUTATIONS	600	CENTROID	336	GROUP	193
HASH	2049	INTERNETWORK	600	DTE	336	PROVISIONED	191
COMPUTATION	2030	DOWNSTREAM	598	OFDM	336	REL	191
BUFFER	2015	PREPAID	597	SVM	336	DEMODULATION	191
ANTENNAS	2010	GPS	596	SUBNETTING	336	CORRELATOR	190
HOST	1997	IGRP	596	LITERALS	335	STATIONARY	190
HOST	1997	PBX	595	MODELLING	334	DECODERS	188
ESTIMATION	1970	STREAMS	591	SUBSYSTEMS	333	INIT	188
HTTP	1947	WEP	590	CONNECTIONLESS	333	SUBPROGRAM	188
FLOW	1928	PACKAGE	588	INGRESS	332	HANDOVERS	188
BITS	1922	BS	582	CLUSTERING	331	HOPPING	188
CONFIGURED	1912	RMI	579	METERS	331	INTERFERING	187
TOOL	1907	ORTHOGONAL	578	MUX	329	EVAL	185
LABVIEW	1882	PROVISIONING	576	ROKE	329	INDUCTORS	185
OSI	1859	DEPENDENT	569	TELEWORKERS	328	DATAPATH	185
FORWARDING	1852	DATAGRAM	568	PSAP	326	SCALEA	182
CONNECTIVITY	1846	FIREWALLS	566	PAC	326	SCHEMAS	181
RADAR	1840	STATIONS	561	VCC	325	NUMBERS	180
PC	1828	PAYLOAD	560	NORMALIZED	324	PDUS	179
CPU	1821	JITTER	559	INTEGRATOR	323	ASICS	179
OPTIMIZATION	1817	MINIMIZATION	559	COAXIAL	322	DEMUX	179
PATHS	1794	OPTIMIZE	555	SSL	322	MATHWORKS	179
GENERATED	1767	DHCP	555	MULTIPROCESSOR	322	DOWNTIME	178
PROPAGATION	1761	WIRE	554	FIFO	321	INTEGRATORS	178
TRANSISTOR	1743	TELEWORK	553	SIM	321	CPUS	177
MICROWAVE	1732	CONFIGURING	553	ESTIMATORS	320	ESERVER	176
MATRIX	1721	ALCATEL	551	CN	319	SDLC	176
PHONES	1715	PHOTONIC	549	HP	319	DIALER	176
SIGNALING	1705	DEPENDING	548	AMS	319	DESKTOPS	175

GENERATION	1697	LSPS	546	HANDSETS	318	GLOBALLY	172
IMPEDANCE	1688	INVARIANT	545	SISO	317	USERNAME	172
COEFFICIENTS	1684	UML	545	PERMITTIVITY	316	HASHING	172
SPECTRAL	1666	WIDEBAND	543	GIS	315	SOLVER	171
ELECTRON	1664	RAM	538	WEBSense	314	SUBCARRIER	169
TABLE	1643	ENCRYPTED	537	OPTIMIZATIONS	314	ROAMABOUT	169
SYNCHRONIZATION	1633	OPTIMISATION	536	VLANS	314	MULTISTREAMING	169
ID	1630	LSP	536	ENDPOINT	313	ORTHOGONALITY	169
SIMULATIONS	1628	ELITECONNECT	536	ASP	312	EQUALIZERS	169
DC	1609	DUPLEX	531	SPEEDUP	312	REACHABILITY	169
SUBSCRIBER	1596	ISP	527	STOPBAND	311	SHIFTER	169
MAC	1588	DETERMININISTIC	527	LIGHTPATH	311	ACTUATORS	169
BAND	1588	ICT	524	GROUPLLET	311	PHOTONS	166
SENDER	1577	STUB	521	CONVOLUTIONAL	311	WIRELINE	166
CMOS	1575	BUFFERS	518	XDSL	311	CONFIG	166
EQUATION	1572	IPX	518	ANALYZER	311	DEMODULATOR	166
GROUPWARE	1571	PDU	518	RECURSION	310	INTERNETWORKS	166
TRANSISTORS	1547	EXPONENTIAL	517	DISTANCES	309	WEBSITE	166
FILTERING	1541	EMITTER	516	PCB	309	EXTENSIBILITY	166
STATIC	1531	CPE	514	HLR	307	TIMEOUTS	166
SIZE	1526	DIPOLE	514	SHWP	307	INDEPENDENTLY	166
WDM	1521	HEADERS	512	INTERNATIONALIZATION	306	LATENCIES	163
SIMULATOR	1508	PHOTONICS	512	RAYLEIGH	306	GROUPLETS	163
WAN	1507	MIB	511	TIMEOUT	306	HSPICE	163
BLUETOOTH	1506	ARP	511	CMT	306	DIALUP	163
CLUSTER	1497	UNICAST	511	GBPS	304	MLINK	163
MBPS	1483	COMBINATIONAL	511	NGSS	304	ALIAS	163
NM	1483	DD	511	NYQUIST	304	IONOSPHERIC	162
UNIX	1481	DISSIPATION	508	MOSFET	303	PAYEE	162
EMBEDDED	1467	CABLES	508	RTS	303	AUTOCORRELATION	160
XILINX	1455	TYPICAL	507	SHELL	303	OUTPERFORMS	160
DIODE	1448	MSC	506	SHELL	303	TOOLSET	160
MAPPING	1435	GENERATES	503	VC	302	NAVSTAR	160
SPACE	1433	DOWNLINK	503	DEPENDS	302	HEURISTICS	160
AMPLITUDE	1430	API	503	DWDM	301	LAPTOPS	159
VPNS	1421	OPTIMUM	502	IFIP	301	METADATA	157
OPTIMAL	1417	DEBUGGING	499	QAM	301	RADIOFREQUENCY	156
WAVE	1408	SCALE	498	CLIENTS	299	DSPS	155
INTEROPERABILITY	1406	CONFIGURABLE	497	SPATIO	298	GENERATORS	155
VOIP	1404	MATRICES	496	YN	298	POINTS	154
LDAP	1374	EXECUTABLE	496	TELELOGIC	298	TELETRAFFIC	154
RECEIVERS	1368	JVM	496	COMPUTATIONALLY	297	MODULATORS	154
MODULES	1356	POLYPHASE	496	MTU	297	MULTIPLEXOR	154
FOURIER	1355	MPEG	494	SYNCHRONIZED	296	RETRANSMITS	154
SUBNET	1355	SUBSYSTEM	490	SIP	296	SUBCLASS	153
WAVELENGTHS	1328	SCHEDULER	490	INTELSAT	295	WAVELETS	152
ATTENUATION	1321	FDDI	489	ADOPTERS	295	PROPAGATES	151
ELECTROMAGNETIC	1304	DOPED	486	BANDWIDTHS	294	METALLIZATION	151
WWW	1303	MICROWAVES	485	PE	294	IMPEDANCES	150
LANS	1302	PDF	484	MODULATOR	293	INTERFACING	150
GPRS	1290	SHARING	483	DDS	292	VOLT	149
OPTIMIZED	1271	SDSL	483	RECONFIGURABLE	292	SPACING	148
TELEPHONY	1266	SMC	483	ENDPOINTS	292	RFEMF	147
ELECTRONICS	1262	NOKIA	481	SLAS	292	LIFECYCLE	147

SETUP	1261	ICMP	480	VBR	292	AMPLITUDES	145
SHARED	1257	PASSBAND	474	RS	292	MULTISERVICE	144
METRICS	1252	UPLINK	474	ACKNOWLEDGMENTS	292	MULTIHOP	144
COMPUTED	1252	TUNNELING	473	ANSI	291	MIBS	144
CARRIERS	1247	MIDDLEWARE	471	VCD	291	URLS	144
CONFIGURE	1240	SPICE	471	RETRANSMIT	290	MULTILAYER	144
SNMP	1239	DLSW	471	PIPELINED	289	EXTRANETS	144
STEP	1239	MN	470	TUNABLE	289	LOOPBACK	144
ASYNCHRONOUS	1225	CONDUCTIVITY	469	VERTEX	289	MBIT	144
BGP	1223	CENTRALIZED	468	DCT	287	PRIORITIZED	144
VOLTAGES	1206	VTP	467	NFS	287	HOSTNAME	141
CONFIGURATIONS	1198	INITIALIZATION	467	SATCOM	285	ASPS	141
ENCRYPTION	1198	DECOMPOSITION	466	REORDERING	285	SUBNETWORKS	141
PAYMENT	1196	LAMBDA	466	ACKS	285	SUBAREA	141
NONLINEAR	1193	BASEBAND	464	PCI	284	RETIMING	141
TELECOMMUNICATION	1174	ABR	464	URL	284	RECURSIVELY	139
SAMPLE	1173	POLYNOMIAL	462	RADARS	283	GLOBALIZATION	138
WAVELET	1169	PLANE	458	VDD	283	GIGAHERTZ	138
GIGABIT	1163	BANDGAP	458	WAVEGUIDES	282	PRIORITIZATION	138
EQ	1158	CSP	458	CROSSTALK	282	LINKSYS	138
GENERATE	1153	DFT	458	CSMA	282	PICOCELL	138
FIREWALL	1147	CELLS	457	LRE	282	METER	137
DCE	1135	FFT	456	TESTBENCH	282	ANALOGOUS	137
SCHEMATIC	1130	SDL	455	REPEATERS	282	SCALED	136
MULTIUSER	1124	GUI	453	PDAS	282	MICROELECTRONICS	136
BIT	1123	MULTIPLIERS	453	DIFFRACTION	281	BANDPASS	135
PROXY	1122	IIR	450	EIFEL	281	ADOPTER	135
SI	1113	ADA	449	OSCILLATORS	281	ENCRYPT	135
SCALABILITY	1112	ITIL	449	GOOGLE	281	SYSTEMC	135
IPSEC	1110	PPTP	449	GB	280	MODELSIM	135
GAUSSIAN	1105	GAM	446	EV	279	MACROMODEL	135
EQUATIONS	1097	INTERCEPTION	445	RCS	279	ETHERWARE	135
MULTIPLEXING	1083	DECNET	445	VERTICES	279	OPAMP	135
MESSAGING	1078	HUB	445	IDS	278	ISERIES	135
ENFORCER	1069	HANDSET	444	LOSSY	278	KTNET	135
MULTIPATH	1069	WAP	441	RMS	276	NONLINEARITIES	135
ITU	1062	URI	441	LOG	276	CONDUCTOR	133
SWITCHED	1059	NETLIST	439	LSRS	276	VERIFIES	132
SUBSTRATE	1047	PSTN	439	DOCSIS	276	BACKPLANES	132
MAGNETIC	1041	HISPASAT	436	PVCS	276	MICROCONTROLLERS	132
HOSTS	1038	RETRANSMISSIONS	436	SMDS	276	UPTIME	132
RFC	1034	MICROCONTROLLER	436	NSA	276	COMBINATORIAL	132
FTP	1021	PREDEFINED	435	WIRING	275	CONDUCTANCE	132
DSL	1020	MPI	434	APPN	275	TEMPORALITY	129
SATELLITES	1020	NARROWBAND	433	TC	274	GAUSS	128

Tabla 10.6 El vocabulario especializado de la ingeniería de telecomunicación.
(Las 1.000 formas más significativas).

CAPÍTULO 11

EL VOCABULARIO ESPECIALIZADO EN LOS TEXTOS ESPECÍFICOS

Al principio de la investigación, se señaló la carencia de material didáctico adecuado para la enseñanza del inglés profesional y académico de la ingeniería en telecomunicación. Los libros de texto existentes estaban anticuados y, a juicio de los docentes, no representaban el contenido léxico de la especialidad en el ámbito académico ni en el profesional. Este problema motivó el presente trabajo.

Una vez que se ha identificado el repertorio léxico especializado tras el análisis empírico realizado sobre la base del corpus específico, se procede primero a comprobar la validez de dicho repertorio y después, a examinar el contenido léxico de los libros de inglés de la especialidad, en relación con el repertorio de vocabulario especializado.

11.1 Validación del repertorio de vocabulario especializado.

El proceso de validación estriba en el filtrado de textos técnicos con el total de las formas que constituyen las diferentes listas de vocabulario: *General Service List* (West, 1953), *Academic Word List* (Coxhead, 2000), el listado de palabras funcionales y el listado de vocabulario especializado de las telecomunicaciones. Todas las formas se unen en una misma lista de corte utilizada para filtrar dos series de textos especializados. En la primera serie se procesan textos pertenecientes al propio corpus técnico de telecomunicaciones y en la segunda, se han empleado textos ajenos al corpus.

La validación del repertorio se argumenta de la siguiente manera: Si las 2.000 familias generales más frecuentes, junto con el vocabulario académico, cubren alrededor del 90% de las palabras de un texto y el 10% restante se divide aproximadamente en partes iguales entre el vocabulario especializado ($\pm 5\%$) y las palabras de baja frecuencia y gran diversidad ($\pm 5\%$) (Laufer, 1989; Hirsh y Nation, 1992; Sutarsyah, Nation y Kennedy, 1994; Nation y Waring, 1997; Nation, 2001), al filtrar un texto con las listas conocidas y el repertorio identificado, las formas restantes deberían ocupar este último 5% del texto.

La primera serie de textos filtrados se ha elegido aleatoriamente entre las muestras recogidas en el corpus, pero intentado abarcar textos de diferente tamaño (el volumen oscila entre 51 y 31.710 *tokens*). El Cuadro 11.1 muestra, para cada texto, el número total de palabras que lo compone, el número de formas que no cubre la lista de corte y el porcentaje correspondiente a este remanente. La incorporación del repertorio especializado a las listas conocidas consigue cubrir una media del 95,6% del texto, es decir, queda descubierto un 4,4% del texto, siendo la tasa máxima de cobertura alcanzada del 98,8% (1,2% sin cubrir) y la mínima del 89% (11% sin cubrir).

Textos filtrados procedentes del corpus

Nº de <i>tokens</i>	Nº de formas restantes	Proporción sin cubrir
51	6	11%
701	43	6,1%
1.314	74	5,6%
1.811	156	8,6%
4.246	153	3,6%
5.030	133	2,4%
5.554	133	2,3%
19.356	222	1,5%
21.968	277	1,2%
31.710	816	2,5%

Cuadro 11.1 Validez de la lista en textos del corpus.

La segunda serie de textos filtrados contiene muestras que no proceden del corpus sino que se han recopilado con el fin de validar el repertorio especializado. Tales muestras cumplen con los requisitos establecidos en el diseño del corpus, datan del año 2005 al 2007 y su tamaño, aun siendo más homogéneo que en el grupo anterior, oscila entre los 323 y 8.963 *tokens*.

El Cuadro 11.2 refleja los datos obtenidos tras el filtrado, donde se observa un leve ascenso en la tasa de cobertura. Ésta alcanza una media del 95,8%, esto es, se queda sin cubrir un 4,2% del texto, obteniendo un máximo del 97,7% (2,3% sin cubrir) y un mínimo del 92,9% (7,1% sin cubrir).

Textos filtrados ajenos al corpus

Nº de <i>tokens</i>	Nº de formas restantes	Proporción sin cubrir
323	23	7,1%
1.841	73	3,9%
2.375	68	2,8%

2.540	169	6,6%
2.822	93	3,3%
3.670	87	2,3%
3.880	122	3,1%
3.981	148	3,7%
5.902	258	4,3%
8.963	473	5,2%

Cuadro 11.2 Validez de la lista en textos ajenos al corpus.

En ambos casos, la fracción de texto que queda descubierta está formada por palabras de diversa índole. Se encuentran unidades léxicas tanto generales como especializadas, nombres propios, palabras con el apóstrofe del genitivo sajón, formas abreviadas, etc... En cuanto a su incidencia, la mayoría son de baja frecuencia y en algunos casos todas las formas presentan frecuencia uno. Por otro lado, en los textos de mayor longitud, hay palabras de frecuencia mucho más alta. Éstas suelen corresponderse con el concepto tratado en el discurso y con otras palabras necesarias para expresar tal concepto. En el siguiente cuadro se ejemplifica con una muestra del corpus y otra externa al corpus la gran diversidad léxica que caracteriza esta pequeña porción de texto.

Muestra del corpus	
Cobertura	<i>Tokens: 7.168</i> Formas restantes: 240 Sin cubrir: 3,3%
Formas de frecuencia alta (39-10) Total: 10	<i>flipper, gizmos, absorber, TA, velocity, bumper(s), jar, gizmo, flipper(s), gizmoball</i>
Formas de frecuencia media (9-5) Total: 18	<i>animation, graphics, gravity, pinball, coefficient, friction, keyboard, rotate, AWT, milestones, triangular</i>
Formas de frecuencia baja (4-2) Total: 50	<i>demo, Goldberg, rotates, clic, settings, collision(s), filterkeys, flip, hardcopy, javax, loop, animate(d)</i>
Formas de frecuencia = 1 Total: 154	<i>palette, additionally, affine, agenda, bounces, bufferedreader, cast, contest, snap, velocities, gameboards, hypotenuse, nuance, SQRT</i>
Muestra ajena al corpus	
Cobertura	<i>Tokens: 3.670</i> Formas restantes: 87

	Sin cubrir: 2,3%
Formas de frecuencia alta (30-10) Total: 10	<i>SRRS, VI, SRM, graph, usage, RMA, VJ, fig.</i>
Formas de frecuencia media (9-5) Total: 6	<i>peer, RI, DS, RJ, dense, subgroup</i>
Formas de frecuencia baja (4-2) Total: 17	<i>competitive, correlated, KN, LU, partition, DAG, spanning</i>
Formas de frecuencia = 1 Total: 49	<i>appeal, candidate, CT, Dijkstra, update, endure, incremental, inter, notation, NI, outperform, session(s), tradeoff, wavy, VK</i>

Cuadro 11.3 Diversidad léxica.

Tras las pruebas realizadas con los diferentes textos, se ha demostrado que la inserción del repertorio especializado en la lista de corte, junto al vocabulario general, funcional y académico, permite alcanzar una tasa de cobertura del texto que llega al 95%. Esta lista deja al descubierto un 5%, en ocasiones una fracción aún menor, correspondiente a una serie de palabras que no se pueden determinar con antelación. En otras palabras, si un lector, el estudiante de ingeniería en telecomunicación en este caso, ampliara su conocimiento del léxico general y académico, con el repertorio de vocabulario especializado, sería capaz de entender el 95% de los textos específicos, encontrando tan sólo una palabra desconocida de cada 20, es decir, alrededor de una palabra por cada dos líneas de texto. Por lo tanto, se puede concluir que el repertorio de vocabulario especializado de las telecomunicaciones es válido.

11.2 El vocabulario especializado en los libros de texto de inglés de la especialidad.

Este apartado da cierre al proceso de investigación realizado, dando respuesta a la última pregunta de investigación planteada: ¿Existen diferencias significativas entre el vocabulario utilizado por la comunidad discursiva de la ingeniería de telecomunicaciones y el vocabulario reflejado en los libros de texto de inglés, destinados a la enseñanza del inglés profesional y académico de la especialidad?

Esta cuestión implica la revisión de los libros de texto de inglés técnico, en concreto, de aquellos disponibles cuando se planteó el desarrollo de un programa para la asignatura de inglés específico. En aquella fecha, no se encontró ningún libro de texto designado concretamente para la enseñanza del inglés de las telecomunicaciones. Los libros más afines a esta ingeniería se centraban en un subdominio constituyente de la especialidad, entre los cuales se optó por los más recientes: “*English for Electronics*” (Álvarez et al., 1990) y “*English for the Telecommunications Engineering Industry*”

(Comfort et al., 1986). Después se incorporó “*Information Technology*” (Glendinning y McEwan, 2002), aun estando más relacionado con la informática.

Los textos recogidos en estos libros se han procesado en WordSmith como se hizo con las muestras del corpus técnico y general, de modo que se puedan obtener datos numéricos que permitan realizar cálculos estadísticos y contrastar los resultados. El tratamiento de los textos se ha realizado tanto de forma individual para cada uno de los libros, como de forma global, considerando todas las muestras en conjunto. La información estadística básica correspondiente en cada caso se indica en el siguiente cuadro:

	Libro 1 Álvarez, 1990	Libro 2 Comfort, 1986	Libro 3 Glendinning, 2002	En conjunto
<i>Tokens</i>	5.383	11.777	11.650	28.810
<i>Types</i>	1.442	2.403	2.418	4.332
<i>Type/Token Ratio</i>	26,79	20,4	20,76	15,04
<i>Standardised Type/Token</i>	42,8	43,98	41,41	42,71

Cuadro 11.4 Información estadística básica de los libros de texto.

Asimismo, se obtienen los cuatro listados de frecuencia pertinentes que posibilitan la comparación del contenido léxico de los libros, en conjunto o por separado, frente al contenido léxico de las telecomunicaciones. Estas listas se utilizan en dos ocasiones, primero para obtener qué palabras son más significativas en los materiales, es decir, para identificar las *keywords*, y después, para comprobar si los libros de texto incluyen las formas inventariadas en el repertorio de vocabulario especializado.

11.2.1 Contraste de listados de palabras clave.

Las palabras clave de los libros se han identificado mediante la comparación de las muestras con el mismo punto de referencia utilizado para el corpus técnico. La presencia de estas palabras clave es estadísticamente significativa respecto a su incidencia en el lenguaje general registrado en el corpus Lacell. Esencialmente, los listados de *keywords* indican qué palabras son más representativas en un área, en este caso, en los libros de texto, al tiempo que muestran cómo difiere y/o se asemeja el comportamiento estadístico de una palabra en los dos discursos comparados. En la Tabla 11.1 se muestra el vocabulario al que cada libro da prioridad, por medio de las primeras 50 palabras clave (el Apéndice X contiene los listados completos):

N	Libro 1	Freq.	Keyness	Libro 2	Freq.	Keyness	Libro 3	Freq.	Keyness
1	CIRCUITS	24	263,5	SUBSCRIBER	31	370,6	DATA	106	684,6
2	CIRCUIT	28	241,9	NETWORK	53	321,4	CACHE	31	396,7
3	TRANSISTORS	15	206,6	TELEPHONE	51	310,3	S	87	354,7
4	COMPONENTS	22	171,5	TRANSMISSION	32	267,5	STORAGE	38	261,3
5	DEVICES	17	133,6	SWITCHING	26	231,6	SOFTWARE	43	259
6	RESISTORS	9	122,6	PRESTEL	18	225,7	MESSAGE	43	247,2
7	INTEGRATED	15	116,2	TELECOM	25	223,4	XML	19	243,6
8	ELECTRONIC	18	116,1	SATELLITES	20	214,5	T	40	180,5
9	PCB	8	101,8	TELECOMMUNICATIONS	23	197,7	USER	29	177,6
10	COMPUTERS	16	101,1	EXCHANGE	35	196,7	LINUX	13	155,1
11	COMPUTER	21	94,1	EXCHANGES	22	192	COMPUTERS	26	149,4
12	DEVICE	13	86,3	DATA	41	188,7	TCP	12	146,7
13	TELEPHONES	9	86	SWEDISH	22	185,2	MESSAGES	20	136,5
14	SATELLITES	8	85,2	PABX	12	179,7	SERVER	19	134,4
15	COMMUNICATIONS	13	82,8	S	53	166	BANDWIDTH	11	124,7
16	NN	5	82,7	SIGNALLING	17	161,7	IP	11	124,7
17	SEMICONDUCTOR	7	78,7	NETWORKS	22	150,9	MAIL	24	122,8
18	USED	32	77,6	DIGITAL	21	140	FRAMES	16	122,1
19	SIGNALS	10	71,9	CIRCUITS	16	138,6	ISP	8	120
20	TELEPHONE	14	71,5	SERVICES	42	138,1	OFFLINE	10	119,7
21	WARRANTY	7	67	EQUIPMENT	29	137,1	LAYER	19	117,8
22	CATHODE	4	66,2	MTX	9	134,8	FRAME	20	117,1
23	DIODES	4	66,2	CCS	11	134,2	PROTOCOL	14	116,5
24	CADAMAT	4	66,2	COMMUNICATION	24	127,3	INTERNET	21	115,2
25	OSCILLOSCOPE	4	66,2	SUBSCRIBERS	14	127,3	HTML	12	110,2
26	COLOR	9	60,8	TRAFFIC	26	124,1	DEVICES	17	107,6
27	ELECTRONICS	8	60,8	VIDEOTEX	8	119,8	HSM	7	105
28	SCHEMATIC	5	59,7	LAN	10	106,6	BLUETOOTH	7	105
29	INSTRUMENT	9	58,7	LANS	7	104,8	SYSTEM	40	98,8
30	RESISTOR	5	58,6	STROWGER	7	104,8	CAN	96	93,8
31	CAPACITORS	5	57,3	INTELSAT	7	104,8	WEB	17	90,9
32	AUTOMATION	6	56,8	SYSTEMS	28	104,7	NETWORK	21	90,4
33	WIRE	9	55,1	MILLION	34	102,8	SGML	6	90
34	TRANSISTOR	5	54,3	CALL	33	95,1	CHECKSUM	6	90
35	PRINTED	9	53,4	TELEX	10	94,3	SMTP	6	90
36	PROCESSING	9	53,1	RADIO	24	91,7	ISPS	6	90
37	DISCRETE	6	52,3	KILOMETRES	13	87,4	ENCRYPTED	6	90
38	MICROCOMPUTERS	5	51,8	COMMUNICATIONS	16	83,8	SYSTEMS	25	88,7
39	ARE	76	51,3	SATELLITE	13	80,7	EMAIL	11	86,3
40	DIGITAL	8	50,7	NUMBER	40	79,5	DISK	13	85,6
41	TROUBLESHOOTING	3	49,6	SYSTEM	35	77,7	LL	9	83,1
42	MINICOMPUTERS	3	49,6	PABXS	5	74,9	KERNEL	8	82,9
43	ICS	3	49,6	COAXIAL	5	74,9	DIGITAL	14	82,7
44	SOLDERED	4	48,8	CABLE	13	74	COCHRANE	8	79,5
45	FUNCTIONS	9	47,5	TELEPHONES	9	71,9	APPLICATIONS	16	78,6
46	MANUFACTURING	8	47,2	STATIONS	13	68,1	UDP	5	75
47	RELIABILITY	6	46,8	SWEDEN	10	67,6	ASP	5	75
48	TOLERANCE	6	46,5	INDIAN	14	66,3	PAYLOAD	5	75
49	TUBES	6	45,9	DIVERSION	9	64,6	MAC	11	72,7
50	SATELLITE	7	45,8	ERICSSON	6	64,1	INFORMATION	35	71,2

Tabla 11.1 Keywords de los libros de texto por separado.

La intersección de las tres muestras y su procesamiento en conjunto provoca cambios en el orden y grado de relevancia de las *keywords*, pero permite obtener una visión más general del vocabulario expuesto en los materiales. De esta forma, la lista de palabras clave (Tabla 11.2) presenta las unidades léxicas relacionadas con las telecomunicaciones que suelen aparecer en los libros, reflejando las diferencias existentes entre el contenido léxico de los libros y el lenguaje general. Estas palabras clave se comparan a su vez con las *keywords* de telecomunicaciones, donde se aprecia una coincidencia del 35% entre las cien palabras más significativas de cada listado, independientemente de la condición académica, general o específica de las formas. Las *keywords* coincidentes se han señalado en negrita en la Tabla 11.2. La muestra que se presenta a continuación contiene las cien primeras palabras clave, encontrándose el documento completo en el Apéndice X:

N	Keywords Teleco	Freq.	Keyness	Keywords Libros	Libros Freq.	Lacell Freq.	Keyness
1	NETWORK	16.649	41.784,60	DATA	158	2.787	860,3
2	DATA	14.613	31.852,20	S	148	7.895	498,1
3	SIGNAL	7.022	17.922,60	NETWORK	81	1.686	415,4
4	SYSTEMS	9.479	17.377,70	TELEPHONE	73	1.606	366,7
5	IP	5.239	16.182,10	CIRCUITS	41	131	352,2
6	NETWORKS	5.832	15.204,90	SUBSCRIBER	34	37	350,2
7	SYSTEM	12.624	14.831,60	CACHE	31	20	340,6
8	USER	6.292	14.725,80	DIGITAL	43	488	269,9
9	PROTOCOL	4.742	13.677,70	SATELLITES	28	52	265,8
10	DESIGN	7.701	12.237,80	TRANSMISSION	38	308	262,5
11	APPLICATIONS	5.414	12.117,70	NETWORKS	38	463	233,4
12	ROUTER	3.910	11.974,40	SOFTWARE	50	1.412	227,5
13	USING	9.214	11.914,70	CIRCUIT	38	525	224,4
14	IS	94.630	11.589,10	SWITCHING	30	187	221,8
15	WIRELESS	4.083	11.454,00	SYSTEMS	62	3.000	219,7
16	FREQUENCY	4.551	11.439,70	COMPUTERS	44	984	219,6
17	FIGURE	7.325	11.299,00	XML	19	12	209,2
18	BASED	8.448	10.804,80	MESSAGE	49	1.628	207,9
19	ROUTING	3.542	10.690,50	DEVICES	35	476	207,7
20	LAYER	4.425	10.604,80	SYSTEM	90	8.707	207
21	MOBILE	4.341	10.529,90	STORAGE	38	803	193,7
22	INPUT	4.347	9.868,60	PRESTEL	18	14	193,6
23	MODEL	5.895	9.860,30	TELECOM	26	177	188
24	INTERNET	4.504	9.651,10	USER	38	903	185,3
25	INTERFACE	3.526	9.557,70	TELECOMMUNICATIONS	26	195	183,3
26	BANDWIDTH	3.119	9.551,10	COMMUNICATIONS	34	779	168,1
27	PACKET	3.577	9.485,10	EXCHANGES	23	174	161,8
28	CIRCUIT	3.932	9.348,50	EXCHANGE	39	1.401	159,8
29	ACCESS	5.999	9.312,50	MESSAGES	28	433	159,3
30	OUTPUT	4.139	9.075,40	PABX	12	0	158,3
31	SERVER	3.574	8.963,20	TRANSISTORS	15	15	156,3
32	USERS	4.493	8.941,90	ELECTRONIC	35	1.036	156,1
33	PERFORMANCE	5.686	8.539,10	T	48	2.879	151,1
34	DIGITAL	3.595	8.516,70	EQUIPMENT	41	1.843	150,8
35	SOFTWARE	4.575	8.470,90	BANDWIDTH	15	20	150,1

36	SIMULATION	2.817	8.189,30	SWEDISH	22	205	146,2
37	DEVICES	3.430	8.086,60	TELEPHONES	19	105	144,7
38	CONTROL	7.124	7.818,40	INTEGRATED	26	444	143,1
39	VOLTAGE	2.945	7.743,20	COMPUTER	49	3.336	143
40	OPTICAL	2.822	7.658,10	SIGNALLING	18	88	141
41	TRAFFIC	4.345	7.420,80	COMPONENTS	28	636	138,9
42	APPLICATION	4.478	7.316,60	LINUX	13	16	131,6
43	ALGORITHM	2.799	7.264,20	INFORMATION	74	9.807	131,1
44	INFORMATION	9.161	7.049,80	SATELLITE	23	384	127,5
45	NODE	2.822	7.042,00	TCP	12	12	125
46	USED	11.874	7.023,70	USED	90	15.508	122,1
47	LINK	3.853	6.946,70	SERVICES	55	5.742	119,3
48	E	6.607	6.856,90	LANS	9	0	118,7
49	SIGNALS	2.893	6.852,30	MTX	9	0	118,7
50	TECHNOLOGY	4.969	6.749,20	CCS	11	11	114,6
51	FILTER	2.627	6.671,30	LAN	12	27	110,2
52	MULTIPLE	3.046	6.621,90	MEMORY	33	1.806	109,4
53	CURRENT	5.064	6.492,40	COMMUNICATION	28	1.130	108,6
54	COMMUNICATIONS	3.144	6.315,60	CAN	178	53.245	107,5
55	CHANNEL	3.212	6.284,10	VIDEOTEX	8	0	105,5
56	FIG	3.702	6.187,30	INTELSAT	8	0	105,5
57	TRANSMISSION	2.544	6.171,40	ISP	8	0	105,5
58	PACKETS	2.308	6.099,60	IP	11	20	104,8
59	NODES	2.361	5.951,60	OPTICAL	16	164	103,5
60	PROCESS	5.949	5.886,10	SUBSCRIBERS	14	91	102,4
61	ROUTERS	1.891	5.876,00	OFFLINE	10	12	101,6
62	WEB	2.978	5.838,20	SERVER	19	362	100,6
63	PROTOCOLS	1.996	5.800,40	LAYER	21	569	97,2
64	VALUE	5.063	5.761,30	RADIO	34	2.460	95,5
65	CIRCUITS	2.122	5.717,70	TRAFFIC	29	1.615	95,2
66	MANAGEMENT	4.890	5.714,70	FRAMES	16	228	93,5
67	SERVICE	7.085	5.707,00	RESISTORS	9	10	92,4
68	DEVICE	2.801	5.665,50	BLUETOOTH	7	0	92,3
69	FIBER	1.869	5.658,80	HSM	7	0	92,3
70	IMPLEMENTATION	2.734	5.617,10	STROWGER	7	0	92,3
71	ETHERNET	1.897	5.601,50	MAIL	26	1.260	92
72	COMPONENTS	2.727	5.589,50	PROTOCOL	14	139	91,4
73	PROCESSING	2.770	5.565,10	DEVICE	21	680	90,1
74	CAN	25.288	5.525,10	HTML	12	75	88,7
75	SERVICES	6.161	5.524,20	FRAME	21	716	88,1
76	EACH	10.224	5.519,00	PROGRAMS	25	1.261	86,6
77	COMMUNICATION	3.159	5.490,80	NUMBER	65	11.407	86,4
78	TCP	1.717	5.248,00	TELEX	11	54	86,1
79	USE	10.255	5.144,70	SIGNALS	17	393	83,8
80	CONFIGURATION	1.885	5.134,10	PROCESSING	20	686	83,7
81	CODE	3.112	5.121,80	PAYLOAD	6	0	79,1
82	ADDRESS	3.951	5.070,10	CHECKSUM	6	0	79,1
83	FUNCTION	3.380	4.960,60	ISPS	6	0	79,1
84	HARDWARE	2.257	4.939,90	COAXIAL	6	0	79,1
85	TECHNOLOGIES	2.137	4.919,50	ENCRYPTED	6	0	79,1
86	ALGORITHMS	1.777	4.828,80	SMTP	6	0	79,1
87	ATM	1.639	4.817,30	SGML	6	0	79,1
88	TYPE	4.612	4.716,70	FREQUENCY	17	455	79,1
89	PROVIDES	3.257	4.553,90	INTERNET	21	910	78,7

90	LAN	1.481	4.387,20	STORED	17	481	77,3
91	RECEIVER	1.699	4.353,70	LL	10	55	76,2
92	DELAY	2.313	4.340,50	MAIN	41	5.177	75,8
93	VALUES	3.008	4.332,70	TECHNOLOGY	31	2.794	75,1
94	ARCHITECTURE	2.581	4.325,80	ERICSSON	8	16	75
95	PARAMETERS	1.841	4.319,50	PCB	8	16	75
96	FRAME	2.327	4.313,30	ARE	276	115.124	72,5
97	FUNCTIONS	2.505	4.252,30	APPLICATIONS	20	934	72,1
98	CONNECTION	2.463	4.222,10	CELLULAR	11	113	71,1
99	MODE	2.128	4.204,20	CALL	41	5.634	70,2
100	PATH	2.700	4.196,70	CABLE	16	501	69,7

Tabla 11.2 *Keywords*

11.3 El repertorio de vocabulario especializado en los libros de texto.

La validación del repertorio de vocabulario especializado permite que éste se establezca como vocabulario de referencia, ya que resulta especialmente útil para comunicarse en el ámbito específico. Consecuentemente, este vocabulario debe aparecer en los libros de texto destinados a la enseñanza del inglés técnico.

A continuación, se procede a comprobar si las palabras del repertorio se encuentran realmente en los libros de texto y en qué medida reflejan el vocabulario especializado. La disponibilidad de los listados de palabras adecuados permite realizar su comparación en una hoja de Excel, como se operó en ocasiones anteriores (apartados 8.1.2 y 8.3). Frente a la lista de vocabulario especializado se antepone el listado de frecuencias de los diferentes libros, de modo que se puede observar qué formas especializadas se incluyen en cada uno (las palabras en negrita se corresponden con las formas más relevantes de cada familia). El Apéndice X contiene el documento completo del cotejo, del cual se presenta una muestra en la Tabla 11.3:

N	Vocabulario especializado	Frec. Teleco	Frec. Lacell	Keyness	Libro 1	Libro 2	Libro 3
1	NETWORK	16.649	1.686	41.784,60	1	2	3
2	DATA	14.613	2.787	31.852,20	1	2	3
3	SIGNAL	7.022	641	17.922,60	1	2	3
4	SYSTEMS	9.479	3.000	17.377,70	1	2	3
5	IP	5.239	20	16.182,10	-	-	3
6	SYSTEM	12.624	8.707	14.831,60	1	2	3
7	PROTOCOL	4.742	139	13.677,70	-	-	3
8	ROUTER	3.910	25	11.974,40	-	-	-
9	WIRELESS	4.083	171	11.454,00	1	-	3
10	FREQUENCY	4.551	455	11.439,70	1	2	3
11	ROUTING	3.542	40	10.690,50	1	-	3
12	LAYER	4.425	569	10.604,80	1	-	3
13	MODEL	5.895	2.290	9.860,30	1	2	3
14	INTERNET	4.504	910	9.651,10	-	-	3
15	INTERFACE	3.526	207	9.557,70	1	2	3
16	BANDWIDTH	3.119	20	9.551,10	1	2	3
17	PACKET	3.577	251	9.485,10	-	2	3
18	CIRCUIT	3.932	525	9.348,50	1	2	-

19	SERVER	3.574	362	8.963,20	-	-	3
20	SOFTWARE	4.575	1.412	8.470,90	1	2	3
21	SIMULATION	2.817	73	8.189,30	-	-	-
22	VOLTAGE	2.945	220	7.743,20	1	-	-
23	OPTICAL	2.822	164	7.658,10	1	2	3
24	TRAFFIC	4.345	1.615	7.420,80	1	2	3
25	ALGORITHM	2.799	229	7.264,20	-	-	3
26	NODE	2.822	294	7.042,00	-	-	3
27	FILTER	2.627	247	6.671,30	-	-	-
28	PACKETS	2.308	166	6.099,60	-	-	3
29	NODES	2.361	232	5.951,60	-	-	3
30	ROUTERS	1.891	4	5.876,00	-	-	-
31	WEB	2.978	791	5.838,20	-	-	3
32	PROTOCOLS	1.996	52	5.800,40	-	-	3
33	CIRCUITS	2.122	131	5.717,70	1	2	3
34	SERVICE	7.085	7.291	5.707,00	1	2	3
35	FIBER	1.869	19	5.658,80	1	-	-
36	ETHERNET	1.897	37	5.601,50	-	-	3
37	SERVICES	6.161	5.742	5.524,20	1	2	3
38	TCP	1.717	12	5.248,00	-	-	3
39	CONFIGURATION	1.885	106	5.134,10	1	-	-
40	HARDWARE	2.257	423	4.939,90	1	2	3
41	ALGORITHMS	1.777	102	4.828,80	-	-	3
42	ATM	1.639	35	4.817,30	-	-	-
43	TYPE	4.612	3.750	4.716,70	1	2	3
44	LAN	1.481	27	4.387,20	1	2	3
45	ARCHITECTURE	2.581	998	4.325,80	-	2	-
46	PATH	2.700	1.210	4.196,70	1	2	3
47	OSPF	1.284	0	4.027,40	-	-	-
48	WAVELENGTH	1.352	24	4.010,00	-	2	-
49	EXAMPLE	6.379	8.009	3.992,60	1	2	3
50	JAVA	1.344	34	3.912,60	-	-	-
51	NUMBER	7.759	11.407	3.799,70	1	2	3
52	SWITCH	2.075	687	3.731,10	-	2	-
53	ANALOG	1.264	27	3.715,00	1	-	-
54	QOS	1.155	0	3.622,70	-	-	-
55	VHDL	1.150	0	3.607,10	-	-	-
56	ANTENNA	1.242	34	3.597,90	-	2	-
57	LINEAR	1.590	276	3.552,30	-	-	3
58	MPLS	1.112	0	3.487,90	-	-	-
59	MODELS	2.255	1.024	3.479,80	-	-	-
60	SOLUTIONS	1.824	523	3.475,60	-	2	3
61	GSM	1.109	4	3.427,30	-	-	3
62	SPECTRUM	1.499	240	3.418,10	-	2	3
63	LINUX	1.128	16	3.377,10	-	-	3
64	INTERFACES	1.187	44	3.364,10	-	2	3
65	POWER	5.031	6.073	3.324,30	1	2	3
66	PROGRAMMING	1.560	326	3.308,90	1	-	3
67	REMOTE	1.770	546	3.277,40	-	2	3
68	CELL	2.089	933	3.252,90	-	-	-
69	CABLE	1.715	501	3.244,60	-	2	3
70	AUTHENTICATION	1.082	16	3.234,20	-	-	3
71	SERVERS	1.208	79	3.231,90	-	-	3
72	VPN	1.007	5	3.097,80	-	-	-

73	FILTERS	1.207	108	3.089,60	-	-	-
74	SWITCHING	1.315	187	3.083,10	1	2	3
75	IEEE	1.002	9	3.044,10	-	-	-
76	BROADBAND	1.049	33	3.010,30	-	-	3
77	SOLUTION	2.392	1.530	2.972,00	1	-	3
78	TOOLS	1.891	832	2.968,30	1	-	3
79	PARALLEL	1.650	569	2.914,90	-	-	-
80	TYPES	2.565	1.850	2.906,10	1	2	3
81	NOISE	1.893	878	2.887,20	-	2	-
82	MEMORY	2.458	1.806	2.743,00	1	2	3
83	SATELLITE	1.401	384	2.715,10	1	2	3
84	LSAS	858	1	2.676,10	-	-	-
85	MODULATION	908	19	2.671,50	-	2	3
86	SWITCHES	1.107	144	2.646,10	1	2	3
87	MULTICAST	837	0	2.625,30	-	-	-
88	DSP	906	41	2.523,20	-	-	-
89	LSA	804	0	2.521,80	-	-	-
90	CDMA	805	1	2.510,00	-	-	-
91	CISCO	840	14	2.498,30	-	-	-
92	MODELING	856	21	2.496,40	-	-	-
93	IMPLEMENTATIONS	826	13	2.462,70	-	-	-
94	RELAY	926	64	2.460,00	1	2	-
95	CELLULAR	987	113	2.420,30	1	2	3
96	DIRECTORY	1.251	348	2.411,10	-	2	3
97	TOPOLOGY	811	14	2.408,40	-	2	-
98	CARRIER	1.102	215	2.384,80	-	2	-
99	SEMICONDUCTOR	845	33	2.384,50	1	-	-
100	GLOBAL	1.800	1.062	2.367,20	1	2	-

Tabla 11.3 Vocabulario especializado en los libros de texto.

A simple vista se puede observar que no todas las palabras aparecen en los tres ejemplares. Además, llama la atención que formas tan relevantes como *protocol*, *internet*, *IP*, *filter*, *routers*, *fibers*, *configuration*, *ATM*, etc... no estén presentes en ningún libro o tan sólo en uno. Los libros comparten 52 palabras, 67 aparecen en dos y 236 en uno, de las cuales 63 pertenecen al libro uno, 58 al dos y 115 al tres.

De las 2.747 formas que constituyen el repertorio de vocabulario especializado, el primer libro incluye 154, el segundo 158 y el tercero 214. Estas cifras se interpretan de varias maneras: en relación con el número de palabras nocionales y el volumen total de cada libro (Figura 11.1) o bien respecto al conjunto de formas del repertorio (Figura 11.2).

La Figura 11.1 representa gráficamente la relación existente entre *tokens*, palabras nocionales y formas especializadas en cada libro. Las muestras de los libros se filtraron con la lista de corte de palabras funcionales, de modo que, los listados de frecuencia resultantes sólo contienen palabras nocionales (Libro 1: 1.315, libro 2: 2.231 y libro 3: 2.253). De éstas, el 11,7%, 7% y 9,4% se corresponden con formas especializadas del repertorio en los libros uno, dos y tres respectivamente. Los porcentajes cambian

sustancialmente al 2,8%, 1,3% y 1,8% calculando la proporción respecto al volumen total de *tokens*. De este modo, teniendo en cuenta la relación entre las diferentes listas de palabras y el fragmento de texto que cubren, el porcentaje de formas especializadas debería aumentar en los libros hasta un 5% aproximadamente, para conseguir una mejor representación del vocabulario.

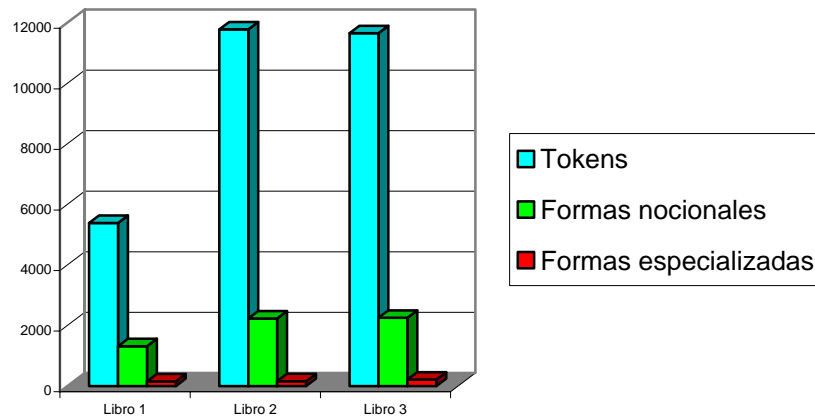


Figura 11.1 Relación formas especializadas – formas conceptuales – *tokens*.

Por otro lado, la Figura 11.2 ilustra la fracción del repertorio de vocabulario especializado que aparece en los libros: un 5,6%, 5,7% y 7,7% para los libros uno, dos y tres respectivamente.



Figura 11.2 Fracción del repertorio de vocabulario especializado en los libros.

El conjunto de datos obtenidos revela que los libros de textos examinados no reflejan adecuadamente el vocabulario especializado, ya que tan sólo incluyen una mínima parte del elenco más significativo. Las carencias detectadas indican que los libros analizados no son lo suficientemente específicos, o bien presentan un vocabulario restringido a una rama de las telecomunicaciones, dejando de lado el resto de sus componentes. Las características léxicas encontradas demuestran que existen diferencias importantes entre el uso real del lenguaje y su representación en los libros de texto, y que por lo tanto, éstos no responden a las necesidades de los alumnos.

CAPÍTULO 12

CONCLUSIONES, APLICACIONES Y FUTUROS TRABAJOS

En el último capítulo de la presente investigación se exponen las principales conclusiones a las que se ha llegado, así como las posibles aplicaciones del estudio, para finalizar con la propuesta de nuevas vías de investigación y futuros trabajos, derivados de los hallazgos y limitaciones de la tesis.

Los resultados obtenidos a lo largo de este estudio, vinculados a los interrogantes planteados y los objetivos perseguidos, e interpretados de acuerdo con el marco teórico establecido, han permitido extraer las siguientes conclusiones:

1. La recopilación de un corpus lingüístico del inglés propio de la ingeniería de telecomunicaciones constituye una de las principales aportaciones de la investigación. El corpus específico es un elemento crucial, puesto que contiene el objeto de análisis, sin el cual no habría sido posible el desarrollo del estudio. El corpus específico registra el uso del lenguaje real y contextualizado de una comunidad discursiva, en la cantidad y variedad suficientes como para ser considerado razonablemente representativo del uso lingüístico.
2. El acopio de muestras almacenadas en formato electrónico ha permitido procesar y cuantificar el lenguaje, para implementar una metodología de análisis propia de la lingüística de corpus, basada en la aplicación de criterios cuantitativos, combinados con criterios cualitativos. La metodología planteada ha posibilitado una aproximación al nivel léxico desde diferentes perspectivas y en función de las variables de frecuencia, distribución, restricción y representatividad de una unidad léxica en el dominio específico.
 - El procesamiento del corpus técnico mediante las listas de corte ha sido válido para revelar qué palabras son comunes a todos los subdominios que integran las telecomunicaciones y cuáles son únicas de cada uno de ellos, otorgando un valor numérico a la distribución de las palabras. Sin embargo, la consideración de las variables de frecuencia y distribución para determinar el carácter especializado de una unidad, no ha sido efectiva. Los resultados obtenidos muestran que, para las diferentes combinaciones de frecuencia

absoluta y distribución, hay palabras especializadas y generales. Por lo tanto, estos parámetros no han identificado comportamientos distintivos de cada categoría.

- La comparación del corpus técnico y un corpus general (corpus Lacell) ha facilitado información esencial sobre el comportamiento de una misma palabra en los dos ámbitos. Al disponer de los valores de frecuencia en cada uno, se han podido realizar cálculos estadísticos de diversa índole que han proporcionado la clave para efectuar una clasificación semiautomática del vocabulario, atendiendo a los diferentes niveles léxicos presentes en un lenguaje de especialidad: vocabulario general, vocabulario técnico y vocabulario semitécnico.

Tras la aplicación de los criterios establecidos por Chung (2003) para la clasificación automática de términos, se ha conseguido distinguir un total de 13.947 formas técnicas. Los criterios de selección tienen en cuenta la frecuencia absoluta de las palabras en los dos corpus y el número de veces que una palabra es más frecuente en el corpus técnico que en general. Esta cifra se sitúa en 50, es decir, para que una palabra se catalogue como especializada debe ser al menos 50 veces más frecuente en el corpus técnico que en el corpus general.

El volumen de términos conseguido comprende tanto palabras técnicas que pueden aparecer en usos no especializados del lenguaje general y palabras generales que adquieren un significado específico, como términos técnicos restringidos al dominio especializado, esto es, el vocabulario técnico con sus diferentes grados de especialización.

- La identificación de las palabras clave ofrece una visión nueva del léxico, donde se valoran las palabras en función del grado de relevancia y representatividad en el sublenguaje analizado. El test estadístico *Log Likelihood* determina que una palabra es clave, cuando su frecuencia en el corpus técnico es significativamente mayor que su frecuencia en el corpus general. El hecho de que las palabras clave se utilicen en el inglés de las telecomunicaciones con mayor frecuencia que en el inglés general, independientemente del nivel léxico al que pertenezcan, les confiere una gran importancia en términos de funcionalidad: cuanto más significativa sea una palabra, mayor será la probabilidad de encontrarla en el lenguaje especializado y por lo tanto, más útil conocer cómo se utiliza.
3. La conjunción de los procedimientos anteriores ha proporcionado la clave para discernir el vocabulario especializado más relevante. La información extraída

del corpus sobre la probabilidad de aparición, la distribución y el comportamiento del léxico especializado constituyen índices de utilidad, cruciales en la selección de contenidos, en la secuenciación de la enseñanza, en la inversión de tiempo y en la importancia atribuidas a las diferentes partes del lenguaje.

- El conocimiento de este vocabulario, junto al académico y al general más frecuente, capacitaría a un lector para comprender el 95% de las palabras presentes en un texto. Diferentes estudios realizados al respecto sostienen que serían necesarias alrededor de 1.000 familias de palabras técnicas para cubrir la proporción correspondiente de vocabulario especializado (5%). Sin embargo, no argumentan cómo distinguirlas.
- El proceso de investigación desarrollado ha discurrido en un método que facilita la identificación semiautomática del vocabulario especializado, capaz de cubrir el porcentaje pertinente en el inglés de las telecomunicaciones.

A la luz de los resultados obtenidos, se ha resuelto que este vocabulario consta no sólo de familias completas, sino también de formas sueltas especializadas, comprendidas entre las 1.000 familias de palabras más significativas del sublenguaje en cuestión. Además, se ha accedido a aquellas formas catalogadas como académicas o generales más frecuentes que adquieren un uso especializado en este registro.

- Algunas formas del vocabulario académico presentan un comportamiento que no es propio de su naturaleza. Las palabras académicas se caracterizan por no estar vinculadas a un dominio en particular, aun siendo frecuentes en una amplia variedad de disciplinas, y no están provistas de significados adicionales. Sin embargo, a juzgar por el comportamiento observado en el corpus técnico, se concluye que tales palabras mantienen su calidad académica siempre que coincidan morfológicamente con las palabras registradas en el listado académico de Coxhead (2000), y además se utilicen individualmente, es decir, cuando no se combinan con otras unidades léxicas, transformando su condición académica en especializada.
 - El volumen del vocabulario especializado alcanza un total de 2.747 formas, entre las cuales se encuentran 402 familias. El número de formas queda considerablemente reducido respecto a la cantidad de palabras identificadas por el método de Chung, siendo a la vez una cantidad más razonable y manejable para fines didácticos.
4. El conjunto de datos proporcionados sobre el comportamiento del léxico respecto a los parámetros de frecuencia, distribución, representatividad y

restricción, se complementa con el análisis detallado, donde se presentan las relaciones semánticas que se establecen entre un lexema y resto que le acompaña en el plano sintagmático.

- La caracterización léxica en el nivel sintagmático se ha llevado a cabo mediante la exploración de los colocados contiguos, los colocados significativos, las combinaciones especializadas y las agrupaciones de palabras o clusters. La determinación de estos rasgos ha demostrado ser decisiva en diferentes ocasiones:
 - o ha corroborado el carácter especializado de las formas identificadas como tal, y por lo tanto, ha validado el método de clasificación;
 - o ha despejado las dudas existentes sobre la valoración automática de algunas formas, de modo que ha permitido la incorporación al vocabulario especializado de formas inicialmente valoradas como no-técnicas, tras comprobar el carácter especializado que adquieren en el contexto;
 - o ha revelado el comportamiento léxico de las formas examinadas, tal y como se utilizan en la comunidad discursiva. La clasificación automática de las unidades léxicas, realizada en función de su morfología, se corrobora o se refuta en el contexto donde se activan los significados, de acuerdo con los rasgos pragmáticos que definen el sublenguaje especializado.

El análisis detallado manifiesta el gran potencial del corpus, respecto a la variedad, la cantidad y la calidad de la información que se puede extraer, ya que sólo se ha presentado una muestra de formas analizadas.

5. Respecto al proceso de validación del repertorio de vocabulario especializado, se ha comprobado que la unión del vocabulario académico, el vocabulario general más frecuente y el repertorio especializado, consigue alcanzar altas tasas de cobertura en textos técnicos. El porcentaje de palabras al descubierto se sitúa en una media del 4,4% para textos procedentes del corpus de Teleco, y del 4,2% para textos ajenos al corpus.

- La validación del repertorio de vocabulario especializado permite que éste se establezca como vocabulario de referencia, dada la importancia de las formas que contiene en términos de representatividad y utilidad, fundamentalmente para la enseñanza del inglés específico. Como consecuencia, los libros de texto deberían incluir este vocabulario. No obstante, los resultados de las acciones realizadas para comprobar en qué

medida los libros reflejan dicho vocabulario, apoyan la impresión inicial sobre la escasa relación entre el contenido léxico de tales libros y el utilizado realmente en un ámbito profesional.

- Partiendo del hecho de que los libros de texto disponibles no están dedicados al estudio del inglés profesional y académico de las telecomunicaciones sino a alguna de sus ramas, y que el contenido léxico que presentan no muestra de forma equilibrada el vocabulario especializado, es necesario exhortar a la elaboración de un libro de texto que cubra las carencias detectadas y atienda a las necesidades reales.
6. El repertorio de vocabulario especializado de la ingeniería de telecomunicaciones comprende las unidades léxicas especializadas que son centrales y típicas en este ámbito. No es un acopio de la terminología específica sino de las formas especializadas, con diferentes grados de especialización, más significativas y representativas, de acuerdo con pruebas estadísticas que cuantifican la probabilidad de aparición de una palabra y su grado de representatividad. Entre las palabras del repertorio se encuentran términos puramente técnicos que no aparecerían en otro dominio, y unidades léxicas propias del lenguaje general, del académico y de otras áreas del saber que adquieren un significado especial en las telecomunicaciones.

De acuerdo con todos los argumentos esgrimidos anteriormente, se aboga en favor del estudio del vocabulario especializado, en el sentido que se le confiere en este trabajo: el vocabulario característico de un sublenguaje, formado por unidades léxicas de diferentes grados de especialización. De este modo, sería posible realizar una aproximación estratégica al vocabulario especializado, considerando primero aquellas formas que son más significativas y cuyo aprendizaje permite un acceso gradual a la complejidad terminológica y conceptual del área específica.

Finalmente, se pretende concluir esta investigación señalando las posibles aplicaciones de los resultados obtenidos, y formulando una serie de propuestas de trabajo que proporcionen una continuidad a la tarea realizada.

La gran ventaja de disponer de un corpus especializado rigurosamente diseñado y recopilado con el fin de representar, en la medida de lo posible, el lenguaje de una comunidad discursiva, permite realizar una amplia variedad de análisis desde diferentes perspectivas.

Siguiendo la línea establecida en la presente investigación, sería posible replicar la metodología empleada para identificar el vocabulario especializado de cualquier otra variedad de la lengua. Asimismo, los estudios contrastivos entre las diferentes

categorías que conforman el corpus serían muy productivos. De este modo, se obtendrían repertorios individuales que reflejan las diferencias y similitudes en el uso del lenguaje dependiendo de la procedencia geográfica, los tipos de texto y las áreas de conocimiento que constituyen la ingeniería de telecomunicaciones.

No obstante, antes de emprender un nuevo cometido, conviene realizar el análisis detallado del resto de formas representativas de las familias de palabras incluidas en el repertorio especializado, con el fin de confirmar su carácter específico e identificar las combinaciones producidas.

La aplicación más inmediata de la identificación del repertorio especializado es la elaboración de un libro de texto para la enseñanza del inglés técnico, adaptado a las necesidades de los ingenieros en formación, que integre el contenido léxico detectado. Además, el listado supone una estimable ayuda para la secuenciación de la enseñanza, ya que el grado de relevancia asociado a cada palabra indica el orden y la cantidad en la que debería aparecer: cuanto más alto sea el índice, antes debería encontrarse y con más frecuencia debería ser utilizada en clase.

Sin embargo, el listado de referencia se limita al uso escrito del lenguaje y un manual didáctico debería incluir la parte correspondiente al lenguaje oral. Por las razones aducidas en el capítulo sobre el diseño y distribución del corpus de Teleco, éste no recoge muestras orales del discurso, carencia que constituye uno de los problemas más apremiantes que resolver.

Este último aspecto está estrechamente relacionado con la actualización del corpus. Tras el transcurso de un determinado espacio de tiempo, es necesario incorporar nuevas muestras con el fin de recoger los cambios que se hayan podido producir en el uso del lenguaje. Igualmente, cuanto mayor sea el volumen del corpus, más fiable será la imagen que se obtenga del lenguaje.

Respecto a los cambios producidos, cabe mencionar la revisión del diseño del corpus en cuanto a las áreas temáticas tomadas como referentes para la recopilación de muestras. Tales áreas se corresponden con los componentes de los planes de estudios cuya renovación es inminente, al tener que adaptarse al Espacio Europeo de Educación Superior. Por lo tanto, se deberá comprobar la validez del corpus y realizar los ajustes pertinentes.

El corpus permanecerá como fuente de información lingüística y base de análisis para la investigación, aunque, por otro lado, se puede trasladar al aula para el uso directo del alumno, quien tendría acceso inmediato a una fuente real y válida del lenguaje, tal y como se utiliza en el entorno profesional y académico que le es propio.

BIBLIOGRAFÍA

- Aguado, G. et al. (2001) *La Investigación en Lenguas Aplicadas: Enfoque Multidisciplinar*. Fundación Gómez-Pardo: UPM.
- Aijmer, K. and Altenberg, B. (Eds.) (1991) *English corpus linguistics: Studies in Honour of Jan Svartvik*. London: Longman.
- Alcaraz Varó, E. y Moody, B. (1983) *Didáctica del inglés: metodología y programación*. Madrid: Editorial Alambra.
- Alcaraz Varó, E. (1990) *Tres paradigmas de la investigación lingüística*. Alcoy: Editorial Marfil.
- Alcaraz Varó, E. (2000) *El inglés profesional y académico*. Madrid: Alianza Editorial.
- Almela et al. (2005) *Frecuencias del español. Diccionario y estudios léxicos y morfológicos*. Madrid: Universitas.
- Almela, M. (2006) *From Words to Lexical Units. A Corpus-Driven Account of Collocation and Idiomatic Patterning in English and English-Spanish*. Berlin: Peter Lang.
- Altenberg B. and Granger S. (2002) *Lexis in contrast: Corpus-based approaches*. Amsterdam: John Benjamins Publishing Company.
- Álvarez, I. et al. (1990) *English for Electronics*. McGraw Hill.
- Anthony, L. (1997) "Preaching to cannibals: A look at academic writing in engineering". In Orr, T. (Ed) *Proceedings: The Japan Conference on English for Specific Purposes 75-86*. Aizuwakamatsu, Japan: University of Aizu.
- Aston, G. (1996) "What corpora for ESP?" [http:// sslmit.unibo.it/gy/pavesi.htm](http://sslmit.unibo.it/gy/pavesi.htm)
- Aston, G. (1997) "Small and large corpora in language learning". In Lewandowska-Tomaszczusk and Melia (Eds.) *Proceedings of the First International Conference on Practical Applications in Language Corpora*.
- Atkins, B. and Zampolli, A. (1994) *Computational approaches to the lexicon*. Oxford: OUP.

- Ball, C. (1996) Tutorial Notes: "Concordances and Corpora". Georgetown University.
- Bahns, J. and Elbaw, M. (1993) "Should We Teach EFL Students Collocations?" *System* 21, 1: 101-114.
- Baker, M. (1988) "Sub-technical Vocabulary and the ESP Teacher. An Analysis of Some Rhetorical Items in Medical Journal Articles". *Reading in a Foreign Language* 4: 91-105.
- Baker, M., Francis, G. and Tognini-Bonelli, E. (Eds.) (1993) *Text and Technology*. In Honour of John Sinclair. Philadelphia/Amsterdam: John Benjamins.
- Baker, P. (2004) "Querying keywords: Questions of Difference, Frequency, and Sense in Keywords Analysis". *Journal of English Linguistics* 32, 4: 346-359.
- Barnbrook, G. (1996) *Language and Computers. A Practical Introduction to the Computer Analysis of Language*. Edinburgh University.
- Barber, C.L. (1962) "Some Measurable Characteristics of Modern Scientific Prose". *Contributions to English Syntax and Philology. Gothenburg Studies in English* 14: 21-43. Stockholm: Almqvist & Wiksell.
- Barrueco et al. (Eds..) (1995) *Lenguas para fines específicos (IV). Investigación y Enseñanza*. Alcalá de Henares: Universidad de Alcalá de Henares.
- Bhatia, V. (1993) *Analysing Genre: Language Use in Professional Settings*. London: Longman.
- Bhatia, V. (2002) "Applied genre analysis: a multi-perspective model" *Ibérica* 4: 3-19.
- Bhatia, V. (2006) "Text and Context in Applied Genre Analysis" Conference of the *American Association for Applied Linguistics (AAAL)*, Montreal, Canada.
- Bauer, L. and Nation, I. S. P. (1993). "Word families." *International Journal of Lexicography* 6, 4: 253-279. Oxford University Press.
- Beaugrande, R. (2000) "Large Corpora, Small Corpora, and the Learning of Language". In Ghadessy, M. et al. (2001) *Small Corpus Studies and ELT*. Amsterdam: John Benjamins.
- Biber, Conrad and Reppen. (1998) *Corpus Linguistics. Investigating Language Structure and Use*. C.U.P.
- Biber, D. and Conrad, S. (2001) "Quantitative Corpus-based Research: Much More Than Bean Counting". *TESOL Quarterly* 35, 2: 331-336.
- Bueno, R. (2003) *Lenguas para Fines Específicos en España través de sus publicaciones (1985-2002)*. Las Rozas: Proyectos Córdon, D.L.

- Butler, C. (2000) "Corpus linguistics and the teaching and learning of languages". En Asta, D. et al. (2000) (Eds.) *Panorama actual de la Lingüística Aplicada. Conocimiento, procesamiento y uso del lenguaje*. 2. Logroño: AESLA.
- Cabré, M.T. (1993) *La terminología. Teoría, metodología, aplicaciones*. Barcelona: Antártida/Empúries.
- Cabré, M.T. (1996) "Lexicología y variación: hacia un modelo integrado" Actas del V *Simposio Iberoamericano de Terminología RITERM*.
- Cabré, M.T. y Vidal, V. (2004) "La combinatoria léxica en la enseñanza y aprendizaje de lenguas para propósitos específicos". En *Las gramáticas y los diccionarios en la enseñanza del español como segunda lengua; deseo y realidad*. Asele, 2004.
- Cabré, M. T. (2004) "Los bancos de conocimiento especializado multilingüe: un nuevo recurso para la traducción". *Actas de "El español, lengua de traducción". II Congreso Internacional*. Toledo: UCLM.
- Cambridge Advanced Learner's Dictionary* (2003) Cambridge: Cambridge University.
- Carbajosa, N. (2003) "El paradigma cognitivo en la enseñanza del inglés profesional y académico". En Durán, P. et al. (Eds.) *Las lenguas para fines específicos y la sociedad del conocimiento*. Universidad Politécnica de Madrid.
- Carbajosa, N. (2002) "A cognitive experience in ESP: teaching vocabulary to Telecommunications students". *ESP-world 2*: 1-14.
- Carlisle, J. (2003) "Fostering Vocabulary Development in Elementary Classrooms". Annual Conference of the Michigan branch of the *International Reading Association*. Grand Rapids, MI.
- Cobb, T. (1997) "Is there any measurable learning from hands-on concordancing?" *System* 25, 3: 301-315. Pergamon Press.
- Comfort, J et al. (1986) *English for the Telecommunications Engineering Industry*. Oxford : OUP.
- Cowie, A. (1998) *Phraseology. Theory, Analysis and Applications*. Oxford: Clarendon Press.
- Coxhead, A. (2000) "A New Academic Word List". *TESOL Quarterly* 34, 2: 213-238.
- Coyne, M. and Edwards, L. (2000). "Big ideas in vocabulary instruction". Paper presentation at the Institute for *Beginning Reading Conference*. Portland, OR.
- Curado Fuentes, A. (2000) "Representativeness and significance factors in ESP texts". *Ibérica* 2: 43-56.

- Curado Fuentes, A. (2001) *A Lexical Common Core in English for Information Science and Technology*. Cáceres: Universidad de Extremadura, Servicio de Publicaciones.
- Curado Fuentes, A. (2002) “Tasks for Business Science and Technology English: Evaluating Corpus-driven Data for ESP”. *ESP-World* 1, 1.
- Curado Fuentes, A. (2003) “El léxico de la gestión empresarial y su tecnología según corpora de inglés específicos”. En Durán, P. et al. (Eds.) *Las lenguas para fines específicos y la sociedad del conocimiento*. Universidad Politécnica de Madrid.
- Chung, T. (2003) “A corpus comparison approach for terminology extraction”. *Terminology* 9, 2.
- Chung, T. and Nation, P. (2003) “Technical vocabulary in specialised texts” *Reading in a Foreign Language* 15, 2.
- Donley, K. and Reppen R. (2001) “Using Corpus Tools to highlight Academic Vocabulary in SCLT”. *TESOL Autumn Quarterly*.
- Doug, A. (2000) “World Wide Web Access to Corpora”. *Cuadernos de Filología Inglesa*. 9, 1: 125-145.
- Dudley-Evans, T. (1997) “An overview of ESP in the 1990s”. In Orr, T. (Ed) *Proceedings: The Japan conference on English for specific purposes*.
- Dudley-Evans, T. and St John, M. (1998) *Developments in English for Specific Purposes*. Cambridge: Cambridge University Press.
- Dudley-Evans, T (2000) “Genre analysis: a key to a theory of ESP?” *Ibérica* 2: 3-11.
- Duque, M. (2000) *Manual de estilo. El arte de escribir en inglés científico-técnico*. Madrid: Paraninfo.
- Dunning, T. (1993) “Accurate Methods for the Statistics of Surprise and Coincidence”. *Computational Linguistics*. 19, 1: 61-74.
- EAGLES (1994). “Corpus Typology: a framework for classification”. Informe Interno N. 2.1 preparado por John M. Sinclair, Universidad de Birmingham, Corpus Linguistics Group.
- EAGLES (1996a). “Text Corpora Working Group Reading Guide”. Documento Eagles (Expert Advisory Group on Language Engineering) EAG-TCWG-FR-2.
- EAGLES (1996b). “Preliminary Recommendations on Corpus Typology”. Documento Eagles (Expert Advisory Group on Language Engineering) EAG-TCWG-CTYP/P.

- Ensslin, A. and Johnson, S. (2007) "Language in the news: investigations into representations of 'Englishness' using 'Wordsmith Tools'". *Corpora: Corpus-based Language Learning, Language Processing and Linguistics* 2, 1: 153-86. <http://www.leeds.ac.uk/linguistics/WPL/WP2006/5.pdf>
- Farrell, P. (1990) "Vocabulary in ESP: a lexical analysis of the English of Electronics and a study of semitechnical vocabulary". *CLCS occasional*; 25. Dublin: Trinity College.
- Feldman, M. (2003) "English Language and Literacy Learning: Research to Practice". *Topics in Adult ESL Education and Family Literacy*. NCLE <http://www.cal.org/caela/ResSLA.htm>.
- Firth, J.R. (1957) "A Synopsis of Linguistic Theory. 1930-1955". In J.R. Firth (Ed.) *Studies in Linguistic Analysis*. Oxford: Basil Blackwell.
- Flowerdew, J. (1991) "Corpus-based Course Design". In J.C. Milton and K.S.T. Tong (Eds.) *Text Analysis in Computer Assisted Language Learning Applications, Qualifications and Developments*. Hong Kong: Language Centre, Hong Kong University of Science and Technology.
- Flowerdew, J. (1993) "Concordancing as a tool in course design". *System* 21, 2: 231-244. Pergamon Press.
- Flowerdew and Peacock (2001) *Research Perspectives on English for Academic Purposes*. Cambridge: Cambridge University Press.
- Flowerdew, L. (2005) "An integration of corpus-based and genre-based approaches to text analysis in EAP/ESP; countering criticisms against corpus-based methodologies." *English for Specific Purposes* 24, 3: 321-333.
- Francis, W.N. and Kucera, H. (1982) *Frequency analysis of the English usage*. Boston: Houghton Mifflin Company.
- Frantzi, K., Ananiadou, S. and Tsujii, J. (1999) "Classifying technical terms". *Redefining the Information Chain - New Ways and Voices*. Proceedings of an ICCC/IFIP conference. Smith, J., Ardö, A and Linde, P. (Eds). Washington D.C.: ICCC Press.
- Gabrielatos, C. (2005) "Corpora and Language Teaching: Just a fling or wedding bells?" *Teaching English as a Second or Foreign Language* 8, 4. www.tesl-ej.org
- García, M. (1994) "Language Needs. for European Engineers. Universidad Politécnica de Madrid: A case study. Need and limits of international cooperation in the fields of higher engineering education in the European context". *EPF* Lausanne, Switzerland.

- García, M. (1995) "Foreign Languages in the Engineering Programs in Spain. The State of the Art". *European Journal of Engineering Education*. Special Issue.
- Gavioli, L. (2005) *Exploring corpora for ESP learning. Studies in corpus Linguistics*. John Benjamins.
- Garside, R. et al. (1987) *The Computational Analysis of English. A corpus-based approach*. London and New York: Longman.
- Granger, S. (1998) *Learner English on computer*. New York: Longman.
- Glendinning, E and McEwan, J. (2002) *Information technology*. Oxford: OUP.
- Hadley, G. "Sensing the Winds of Change: An Introduction to Data-driven Learning". To appear in *Insights 2*.
- Hermans, S. (1997) "Promoting foreign language competence in the European Community: The LINGUA programme." *World Englishes* 16, 1: 45-55.
- Hewings, M. (2002) A History of ESP through 'English for Specific Purposes'. *ESP-World* 3, 1.
- Hoey, M. (1991) *Patterns of lexis in text*. Oxford, OUP.
- Hoey, M. (1997) "From concordance to text structure: new uses for computer corpora". Melia, J. and Lewandoska, B. (Eds.) *Proceedings of PALC 97*. Lodz: Lodz University Press.
- Hoey, M. (2005) *Lexical priming. A new theory of words and language*. London. Routledge.
- Hunston, S. and Francis, G. (1999) *Pattern Grammar: A corpus-driven approach to the lexical grammar of English*. Amsterdam: John Benjamins Publishing Company.
- Hutchison, T. and Waters, A. (1998) *English for Specific Purposes*. Cambridge University Press.
- ISO 704-1987 Principes et méthodes de la terminologie.
- Jackson, H. (1988) *Words and their meaning*. London: Longman.
- Jackson, H. (2002) *Lexicography*. London: Routledge.
- Johansson S. and Knut Hofland (1989) *Frequency analysis of English vocabulary and grammar (based on the LOB Corpus)*. Oxford: OUP.
- Johns, T. (1986) "Micro-Concord: A Language Learner's Research Tool". *System* 14, 2: 151-162.
- Johns, T. (1991) "Should you be persuaded: Two examples of data-driven learning". In Johns, T. and King, P. (Eds.) *Classroom Concordancing*. Birmingham: ELR.

- Johns, T. (1996) "If our descriptions of language are to be accurate . . . A footnote to Kettemann". *TELL&CALL* 4: 44-6.
- Johns, T. (1997) "Kibbitzing One-to-Ones". (Web version of notes for presentation at BALEAP meeting on Academic Writing, University of Reading, 29th November 1997).
- Kennedy, G. (1998) *An introduction to corpus linguistics*. New York: Longman.
- Kennedy, G. y Bolitho, R. (1984) *English for specific purposes*. London: Mcmillan.
- Kilgarriff, A. and Salkie, R. (1996) "Corpus similarity and homogeneity via word frequency". *Euralex 96 Proceedings* Göteborg: 1996.
- Kilgarriff, A. and Grefenstette, G. (Eds.) (2003) Web as corpus. Special issue of *Computational Linguistics* 29, 3. <http://acl.ldc.upenn.edu>
- King, P., Woolls, D. (1996) "Creating and using a multilingual parallel concordancer". *Translation and Meaning* 4: 459-466.
- Leech, G. (1991) "The state of the art in corpus linguistics." In Aijmer, K. and Altenberg, B. (Eds.) *English corpus linguistics: Studies in Honour of Jan Svartvik*. London: Longman.
- Leech, G (1997) Teaching and Language Corpora: a Converge. In Wichman et al. (Eds.) *Teaching and Language Corpora*. Longman.
- Leech, G. et al. (2001) *Word frequencies in written and spoken English*. Longman.
- Lewis, M. (1993) *The Lexical Approach*. England: LTP.
- Lewis, M. (1997) *Implementing the Lexical Approach*. England: LTP.
- Lewis, M. (2000) *Teaching Collocation Further Developments in the Lexical Approach*. England: LTP.
- Lyons, J. (1977) *Semantics*. Vol. I & II. Cambridge University Press.
- Marcos, F. (1994) *Informática y humanidades*. Gredos.
- McEnery, T. and Wilson, A. (1996) *Corpus Linguistics*. Edinburgh: Edinburgh University Press.
- McEnery, T. and Wilson, A. (1993) "The role of corpora in computer-assisted language learning", *Computer Assisted Language Learning* 6, 3: 233-48.
- Meyer, C. (2006) "Corpus Linguistics, the World Wide Web, and English Language Teaching" *Ibérica* 12: 9-21.
- Meyer, I. and Mackintosh, K. (1996) "The Corpus from a Terminographer's Viewpoint". *International Journal of Corpus Linguistics* 1, 2: 257-285.

- Ming-Tzu, K.W. and Paul Nation (2004). "Word Meaning in Academic English: Homography in the Academic Word List" *Applied Linguistics* 25, 3: 291-314.
- Moreno, S y Buyse, K. (2003) "Colocaciones léxicas: pistas y trampas". *Mosaico* 10: 10-17. www.sgci.mec.es/be/MOSAICO10.pdf
- Moore, N. (2002) "Identifying an ESP Vocabulary Syllabus". (Web version)
- Moudraia, O. (2004) "The Student Engineering English Corpus". *ICAME Journal* 28: 139-143.
- Munby, J. (1978) *Communicative Syllabus Design*. Cambridge: Cambridge University Press.
- Mudraya, O. (2005) "Engineering English: A lexical frequency instructional model". *English for Specific Purposes* 25, 2: 235-256.
- Nation, P. (2001) *Learning vocabulary in another language*. Cambridge: Cambridge University Press.
- Nation, P. (1997) "The Language Learning Benefits of Extensive Reading". *The Language Teacher* 21, 5. www.jalt-publications.org
- Nattinger, J. & J. DeCarrico (1992) *Lexical Phrases in Language Teaching*. Oxford University Press.
- Navarro, M. y Brady, I. (2002) "El potencial de los materiales auténticos en el aprendizaje de inglés específico dentro del marco universitario". Actas del II Congreso de la Asociación de Centros de Lenguas en la Enseñanza Superior. Universidad Politécnica de Cartagena.
- Nelson, M. (2000) *A corpus-based study of business English and business English teaching materials*. University of Manchester.
- Oakes, M. P. (1998) *Statistics for Corpus Linguistics*. Edinburgh: Edinburgh University Press.
- Oakey, D. (2004) "Academic Vocabulary in Academic Discourse: the phraseological behaviour of evaluation in the discourse of Economics".
- Palmer, J. et al. (Eds..) (2001) *Discourse Analysis and Terminology in Languages for Specific Purposes*. Universitat Jaume I.
- Palmer, J. et al. (Eds..) (2001) *Methodology and New Technologies in Languages for Specific Purposes*. Universitat Jaume I.
- Pamies, A. y Pazos, J. (2003) "Acceso automatizado a fraseologismos y colocaciones en corpus no etiquetado". *Language Design* 5: 39-50.

- Parodi, G. (2004) "Textos de especialidad y comunidades discursivas técnico-profesionales: una aproximación basada en corpus computerizado". *Estudios Filológicos* 39: 7-36. Chile: Scielo.
- Partington, A. (1998) *Patterns and Meanings. Using Corpora for English Language Research and Teaching*. Amsterdam: John Benjamins Publishing Company.
- Pearson, J. (1998) *Terms in Context*. Amsterdam: John Benjamins Publishing Company.
- Peñas, A. et al. (2001) "Corpus-based terminology extraction applied to information access". 13. Special Issue. Proceedings of the *Corpus Linguistics 2001*. UCREL Technical Paper. Lancaster University. <http://elies.rediris.es/elies18>
- Pérez, Ch. (2002) "Explotación de los corpórea textuales informatizados para la creación de bases de datos terminológicas basadas en el conocimiento". *Estudios de Lingüística Española*, 18. Universidad de Málaga.
- Pikulski, J. and Templeton, S. (2004) "Teaching and Developing Vocabulary: Key to Long-Term Reading Success". www.eduplace.com
- Quereda, L. (1992) "El modelo lingüístico de J.R. Firth" en Quereda L. y Santana J. (Eds.), *Homenaje a J.R. Firth en su centenario*. Servicio Publicaciones Universidad de Granada.
- Rea, C. y Carrillo, M. (2002) "Corpus lingüístico del inglés de Ingeniería de Telecomunicaciones: diseño, recopilación y expectativas" *Actas del II Congreso de la Asociación de Centros de Lenguas en la Enseñanza Superior*. Universidad Politécnica de Cartagena.
- Robinson, P. (1991) *ESP Today: A Practitioner's Guide*. Hertfordshire: Prentice may.
- Rondeau, G. (1983). *Introduction à la terminologie*. Québec: Gaëtan Morin Éditeur.
- Sager, Dungwort and McDonald (1980) *English Special languages. Principles and practice in science and technology*. Wiesbaden, Brandstetter Verlag KG.
- Sager, J. (1990) *A Practical Course in Terminology Processing*. Amsterdam/Philadelphia: John Benjamins.
- Sánchez, A. (2000) "Language Teaching Before and After "Digitalized Corpora". Three Main Issues". *Cuadernos de Filología Inglesa*. 9, 1: 5-37.
- Sánchez, A., Cantos, P. (1998) El ritmo incremental de palabras nuevas en los repertorios de textos. Estudio experimental y comparativo basado en dos corpus

- lingüísticos equivalentes de cuatro millones de palabras, de las lenguas inglesa y española y en cinco autores de ambas lenguas. *Atlantis* 1, 2: 205-227.
- Sánchez, A., Cantos, P., Sarmiento R., Simón, J. (1995) *Cumbre. Corpus lingüístico del español contemporáneo. Fundamentos, metodología y análisis*. Madrid: SGEL.
- Sánchez, P. y Aguado, P. (2003) “El conocimiento previo del vocabulario semitécnico como herramienta útil para el diseño de un curso de inglés para fines específicos académicos (ESAP)”. En Durán, P. et al. (Eds.) *Las lenguas para fines específicos y la sociedad del conocimiento*. Universidad Politécnica de Madrid.
- Sanz, I. y Felices, A. (Eds.) (2004) *Las nuevas tendencias de las lenguas de especialidad en un contexto internacional y multicultural*. Universidad de Granada.
- Schmitt, N. (2000) *Vocabulary in Language Teaching*. Cambridge: Cambridge University Press.
- Scott, M. (1997) “PC analysis of key words – and key key words.”, *System* 25, 2: 233–245. Pergamon Press.
- Scott, M. (1998) *WordSmith Tools Manual version 3.0*. Oxford University Press.
- Sinclair, J. (1991) *Corpus, Concordance and Collocation*. Oxford: Oxford University Press.
- Sinclair, J. (1987) *Looking up. An account of the Cobuild Project in lexical computing*. London: Collins.
- Sinclair, J. (2004) “Intuition and annotation - the discussion continues”. *Advances in corpus linguistics. Papers from the 23rd International Conference on English Language Research on Computerized Corpora (ICAME 23)*. Aijmer, K and Altenberg, B. (Eds.) Amsterdam/New York: Rodopi.
- Stevens, V. (1995) “Concordancing with Language Learners: Why? When? What?” *CAELL Journal* 6, 2: 2-10.
- Stubbs, M. (1986) “Language development, lexical competence and nuclear vocabulary”. *Educational Linguistics* 98-115. Blackwell.
- Stubbs, M. (1993) “British Traditions in Text Analysis: From Firth to Sinclair”. In Baker, M., Francis, G. and Tognini-Bonelli, E. (Eds.) *Text and Technology. In Honour of John Sinclair*. Philadelphia/Amsterdam: John Benjamins.
- Stubbs, M. (1996) *Text and Corpus Analysis*. Oxford: Blackwell.

- Stubbs, M. (2004) "Which words must students know?: a note on "A New Academic Word List" (Coxhead 2000)" www.uni-trier.de/uni/fb2/anglistik/Projekte/stubbs
- Stubbs, M. (2001) *Word and phrases. Corpus studies of lexical semantics*. Oxford: Blackwell.
- Swales, J. (1985) *Episodes in ESP*. Hemel: Hempstead Prentice Hall International.
- Swales, J. (2000) "Language for specific purposes". *Annual Review of Applied Linguistics*, 20: 59-76.
- Swales, J.M. (2004) "Then and now: A reconsideration of the first corpus of scientific English". *Ibérica* 8: 5-21.
- Teubert, W. (1999) "Corpus Linguistics – A Partisan View". *International Journal of Corpus Linguistics* 4, 1: 1-16.
- Thomas, J. and Short, M. (Eds.) (1996) *Using corpora for Language Research. Studies in the Honour of Geoffrey Leech*. London: Longman.
- Tognini-Bonelli, E. 2001. *Corpus linguistics at work: Studies in corpus linguistics*. Amsterdam: John Benjamins.
- Torruella, J. y Llisterri, J. (1999) "Diseño de corpus textuales y orales". En Blecua, J.M., Clavería, G., Sánchez, C. y Torruella, J. (Eds.) *Filología e Informática. Nuevas tecnologías en los estudios filológicos*, 45-77. Barcelona: Editorial Milenio.
- Tribble, C. and Jones, G. (1990) *Concordances in the Classroom*. London: Longman.
- Tribble, C. (1997) "Improvising Corpora for ELT: Quick and Dirty Ways of Developing Corpora for Language Teaching". In Lewandowska-Tomaszczyk, B. and Melia, J. (Eds.) *Proceedings of the First International conference on Practical Applications in Language Corpora*.
- Tribble, C. (2000) "Genres, Keywords, Teaching: towards a pedagogic account of the language of Project Proposals". In Burnard, L. and McEnery, T. (Eds.). *Rethinking language pedagogy from a corpus perspective: papers from the third international conference on teaching and language corpora*. (Lodz Studies in Language). Hamburg: Peter Lang.
- Vangenhuchten, L. (2003) "Algunos resultados de un análisis terminológico de los términos compuestos en un corpus de discurso económico empresarial". En Durán, P. et al. (Eds.) *Las lenguas para fines específicos y la sociedad del conocimiento*. Universidad Politécnica de Madrid.
- Varantola, K. (1986) "Special Language and General Language: Linguistic and Didactic Aspects". In *ALSED-LSP Newsletter* 10, 2, 1987.

- Viel, J.C. (1999) *Science Teaching Reconsidered, A Handbook*.
<http://www.nap.edu/readingroom/books/str/7.html>
- Widdowson, H.G. (1993) "Communication, Community and the Problem of Appropriate Use". In J.E. Alatis (Ed.) *Language, Communication and Social Meaning*. Washington D.C.: Georgetown University Press.
- Wang, K. and Nation, P. (2004) "Word Meaning in Academic English: Homography in the Academic Word List". *Applied Linguistics* 25, 3: 291-314.
- Waring, R. and Nation, P. (1997) "Vocabulary size, text coverage and word lists". In Schmitt, N. and McCarthy, M. (Eds.) *Vocabulary: Description, Acquisition and Pedagogy*. Cambridge University Press.
- Webster's Online Dictionary*. <http://www.websters-online-dictionary.com>
- West, M (1953) *A General Service List of English Words*. London: Longman.
- Wichmann, A. et al. (1997) *Teaching and Language Corpora*. New York: Longman.
- Widdowson, H. (1998) "Communication and community: The pragmatics of ESP". *English for Specific Purposes* 17,1: 3-14.
- Wynne, M. (Eds.) (2005) *Developing Linguistic Corpora: a Guide to Good Practice*. ASDS Literature, Languages and Linguistics. Oxford.
- Yang Huizhong (1986) "A New Technique for identifying Scientific/Technical Terms and Describing Science Texts (An Interim Report)". *Literary and Linguistic Computing*, 1, 2: 93-103.