

## La definición terminológica en vocabularios normalizados UNE: primera exploración de los datos

Seminario IULATerm  
Tesis doctoral dirigida por: Dr. Carles Tebé y Dra Guadalupe Aguado  
Universitat Pompeu Fabra  
16 de marzo de 2010

### Esquema de la presentación

1. **Presentación de los ejes vertebradores de la tesis**
  - 1.0 Introducción
  - 1.1 Delimitación del campo de estudio
  - 1.2 Hipótesis y objetivos del trabajo
  - 1.3 Corpus de análisis del trabajo
2. **Primera exploración de los datos**
  - 2.1 Material analizado
  - 2.2 Construcción de una base de datos provisional
  - 2.3 Identificación de los fenómenos de no correspondencia más significativos
3. **Presentación de los fenómenos y pistas metodológicas para la extracción de las UT y rasgos conceptuales identificados**
  - 3.1 No correspondencia formal entre UT vocabulario-texto
  - 3.2 No correspondencia formal entre UT texto-vocabulario
  - 3.3 Contenido conceptual diferente entre UT texto-vocabulario
4. **Objetivos a corto plazo**

1. Presentación de los ejes vertebradores de la tesis
- 1.0 Introducción

- Tiene como origen el **Trabajo Final de Máster (2007) *Criterios de elaboración de definiciones terminológicas en normas UNE***.
- Ha tenido continuidad en el **Proyecto de Tesis (2009) *La definición terminológica en vocabularios normalizados UNE: análisis y propuesta metodológica***.
- Se enmarca en el proyecto de investigación **TEXTERM 3: *Fundamentos, estrategias y herramientas para el procesamiento, extracción y representación de información especializada*** [HUM2006-09458] financiado por el Ministerio de Educación y Ciencia.

1. Presentación de los ejes vertebradores de la tesis
- 1.1 Delimitación del campo de estudio

- El **campo de estudio** de este trabajo es la normalización y más concretamente los documentos producidos por el organismo español de normalización AENOR:
  - las **normas textuales UNE** fijan prescripciones y requisitos, y establecen las buenas prácticas de un ámbito técnico dado;
  - los **vocabularios normalizados UNE** que fijan las unidades terminológicas utilizadas en las normas textuales.
- El **objeto de estudio** es la representación conceptual en vocabularios normalizados UNE.

1. Presentación de los ejes vertebradores de la tesis
- 1.2 Hipótesis, objetivos del trabajo, y marco teórico

- **Hipótesis general:** Una cantidad significativa de unidades terminológicas procedentes de la **terminología *in vitro*** –los vocabularios normalizados UNE– no corresponden ni formal ni conceptualmente con las unidades terminológicas procedentes de la **terminología *in vivo*** – los textos especializados, en este caso, las normas técnicas UNE (conceptos prestados de Cabré, 1999).

1. Presentación de los ejes vertebradores de la tesis
- 1.2 Hipótesis, objetivos del trabajo, y marco teórico

- **Hipótesis específicas:**
  - Existe un **bajo grado de correspondencia** entre las UT de las normas textuales y las normas de vocabulario;
  - Existe un **bajo grado de correspondencia** entre las variantes denominativas de la norma textual con respecto a la norma de vocabulario;
  - Los textos de las normas técnicas UNE contienen **rasgos conceptuales relevantes** ausentes en las definiciones de las normas de vocabulario temáticamente correspondientes.

La definición terminológica en vocabularios normalizados UNE: primera exploración de los datos

1. Presentación de los ejes vertebradores de la tesis  
1.2 Hipótesis, objetivos del trabajo, y marco teórico

• **Premisa de investigación:**

El análisis de la información terminológica (tanto denominativa como conceptual) extraída de corpus textuales es el procedimiento más adecuado para arrojar luz sobre el funcionamiento del léxico.

"Standardization only seeks users in achieving a greater than normal degree of economy, precision and appropriateness in their communication, i.e. it only aims at maximizing tendencies which are already present in general language." (Sager, 1990)

**iulaTerm** INSTITUT UNIVERSITARI DE LINGÜÍSTICA APLICADA  
UNIVERSITAT POMPEU FABRA

La definición terminológica en vocabularios normalizados UNE: primera exploración de los datos

1. Presentación de los ejes vertebradores de la tesis  
1.2 Hipótesis, objetivos del trabajo y marco teórico

• **Objetivos y marco teórico del trabajo:**

- Presentar una **propuesta de elaboración de definiciones** para vocabularios normalizados en normas UNE a partir de los principios teóricos de la **Teoría Comunicativa de la Terminología** y de los principios metodológicos de la **Lingüística de corpus**.
- Establecer una **metodología de trabajo** que, partiendo del trabajo en corpus, proporcione al redactor terminográfico unas bases sólidas para acometer su trabajo redaccional y permita lograr una mayor armonía terminológica entre vocabularios y normas técnicas UNE.
- Diseñar una **herramienta de ayuda a la elaboración de definiciones**.

**iulaTerm** INSTITUT UNIVERSITARI DE LINGÜÍSTICA APLICADA  
UNIVERSITAT POMPEU FABRA

La definición terminológica en vocabularios normalizados UNE: primera exploración de los datos

1. Presentación de los ejes vertebradores de la tesis  
1.3 Corpus de estudio del trabajo

**Metodología de trabajo**

- Confrontación de un corpus terminográfico con un corpus de normas textuales.
- Correspondencia temática y de fecha entre ambos corpus.
- Área temática elegida: Protección de la salud.

**iulaTerm** INSTITUT UNIVERSITARI DE LINGÜÍSTICA APLICADA  
UNIVERSITAT POMPEU FABRA

La definición terminológica en vocabularios normalizados UNE: primera exploración de los datos

1. Presentación de los ejes vertebradores de la tesis  
1.3 Corpus de estudio del trabajo

|                                  | Bloque 1   | Bloque 2   | Bloque 3   |
|----------------------------------|--|--|--|
| Subcorpus de vocabularios UNE    | UNE-EN 132: 1999 Equipos de protección respiratoria. Definiciones y pictogramas.   | UNE-EN 165: 2006 Protección individual de los ojos. Vocabulario  | UNE-EN-614: 2006 Seguridad de las máquinas. Principios de diseño ergonómico. Parte 1: Terminología y principios generales. |
| Subcorpus de normas de texto UNE | UNE-EN 137: 2007<br>UNE-EN 138: 1995<br>UNE-EN 140: 1998<br>UNE-EN 142: 2002<br>UNE-EN 143: 1997<br>UNE-EN 145: 1998<br>UNE-EN 149: 2001<br>UNE-EN 250: 2001<br>UNE-EN 268: 1995<br>UNE-EN 12941: 1999 | UNE-EN 166: 2002<br>UNE-EN 167: 2002<br>UNE-EN 168: 2002<br>UNE-EN 172: 1995<br>UNE-EN 207: 1999<br>UNE-EN 208: 1998<br>UNE-EN 379: 2004 | UNE-EN 542-2: 1997<br>UNE-EN 894-1: 1997<br>UNE-EN 894-2: 1997<br>UNE-EN 894-3: 1997                                       |

**iulaTerm** INSTITUT UNIVERSITARI DE LINGÜÍSTICA APLICADA  
UNIVERSITAT POMPEU FABRA

La definición terminológica en vocabularios normalizados UNE: primera exploración de los datos

2. Primera exploración de los datos  
2.1 Material analizado

- Siguiendo la metodología de investigación del trabajo de tesis:

**UNE-EN 132 (corpus terminográfico) Vs. UNE-EN 137 y UNE-EN138 (corpus textual)**

- **Procesamiento y análisis manual** de los archivos de texto (adquiridos hace 15 días) en formato desprotegido tras una reunión de colaboración con AENOR;
- **Almacenamiento de los datos** en una base de datos;
- **90 UT** analizadas en total.

**iulaTerm** INSTITUT UNIVERSITARI DE LINGÜÍSTICA APLICADA  
UNIVERSITAT POMPEU FABRA

La definición terminológica en vocabularios normalizados UNE: primera exploración de los datos

2. Primera exploración de los datos  
2.2 Construcción de una base de datos provisional

- El objetivo de la base de datos provisional consiste en **almacenar toda la información relativa a esta primera exploración** de los datos para:
  1. Visualizar con claridad los fenómenos más significativos que resultan del contraste del vocabulario normalizado con las normas de texto;
  2. Restringir el trabajo de tesis al estudio de uno o algunos de estos fenómenos;
  3. Tomar las decisiones adecuadas para el diseño de la base de datos definitiva.
- La base de datos es **relacional** y consta de 5 tablas:
  - BD general;
  - UT frecuentes;
  - UT coincidentes;
  - Variantes;
  - Análisis verbos.

**iulaTerm** INSTITUT UNIVERSITARI DE LINGÜÍSTICA APLICADA  
UNIVERSITAT POMPEU FABRA

2. Primera exploración de los datos  
2.3 Identificación de los fenómenos de no correspondencia más significativos

- Los fenómenos más significativos que resultan del contraste de la norma de vocabulario UNE-EN 132 con las normas técnicas UNE-EN 137 y UNE-EN 138 son los siguientes:

- No correspondencia formal entre UT vocabulario-texto;
- No correspondencia entre las variantes de vocabulario y texto;
- Contenido conceptual diferente entre UT texto-vocabulario.

3. Presentación de los fenómenos y propuesta metodológicas para la extracción de las UT y rasgos conceptuales identificados

3.1 No correspondencia formal entre UT vocabulario-texto

- **Supuesto previo:** todas las UT relevantes de una norma de texto deberían estar recogidas en la norma de vocabulario correspondiente.
- **Ejemplo:** De las **90 UT** extraídas de la norma de texto UNE-EN 137, sólo **11** están recogidas en la norma de vocabulario UNE-EN 132.

**UT relevantes no recogidas:** botella a presión con válvula, presión máxima permitida, mangueras, conector de media presión, equipo de respiración autónomo de circuito abierto de aire comprimido, válvula de cierre, botella a presión, botella de aire comprimido, dispositivo de estanquidad, conexión permanente, conexión a rosca, etc.

3. Presentación de los fenómenos y propuesta metodológicas para la extracción de las UT y rasgos conceptuales identificados  
3.1 No correspondencia entre las variantes de vocabulario y texto

- **Pregunta de investigación:** ¿Cómo detectar las UT relevantes?
- **Estrategias a utilizar:**
  1. Utilización de patrones morfosintácticos;
  2. Utilización de marcadores estructurales: las UT relevantes pueden aparecer en partes relevantes del texto (títulos, subtítulos, apartados de *Requisitos*, etc.)
  3. Utilización de marcadores lingüísticos: las UT relevantes inmediatamente antes de verbos de función como *deber* o *poder*.
  4. Detección de la frecuencia de aparición: las UT frecuentes en un texto son generalmente UT relevantes del ámbito temático tratado.
  5. Combinación de las distintas estrategias.

3. Presentación de los fenómenos y propuesta metodológicas para la extracción de las UT y rasgos conceptuales identificados  
3.1 No correspondencia entre las variantes de vocabulario y texto

- **Metodología:**
  1. Lista de patrones morfo-sintácticos;
  2. Marcaje estructural del texto;
  3. Marcaje lingüístico del texto;
  4. Extracción de concordancias;
  5. Obtención de lista de candidatos a UT que deberían figurar en la norma de vocabulario;
  6. Validación por expertos de AENOR.

3. Presentación de los fenómenos y propuesta metodológicas para la extracción de las UT y rasgos conceptuales identificados  
3.2 No correspondencia entre las variantes de vocabulario y texto

- **Supuesto previo:**
  - las variantes de las normas de texto deberían estar recogidas en las normas de vocabulario correspondiente;
  - las variantes denominativas (texto-texto; vocabulario-texto) conllevan variación conceptual.

3. Presentación de los fenómenos y propuesta metodológicas para la extracción de las UT y rasgos conceptuales identificados  
3.2 No correspondencia entre las variantes de vocabulario y texto

• **Ejemplos:**

| Variante en UNE-EN132   | Variante en UNE-EN 137  | Fenómeno observado   |
|---|---|--|
| arnés   | arnés para el cuerpo  | <b>Variación denominativa:</b><br>un rasgo de la definición pasa a formar parte de la variante en la norma de texto.   |
| media presión   | presión media   | <b>Variación denominativa:</b><br>los compuestos de la UT están invertidos en vocabulario y texto.   |
| indicador de presión<br>(hiperonimo de indicador de presión tipo aguja, indicador de presión tipo táctil) | indicador de presión<br>(sinónimo de indicador de presión tipo aguja, indicador de presión tipo táctil) | <b>Variación conceptual:</b><br>En la norma de texto, la UT correspondiente al concepto genérico y las UT correspondientes a reducciones analíticas (Freixá, 2012) de conceptos específicos coinciden. |

La definición terminológica en vocabularios normalizados UNE: primera exploración de los datos

3. Presentación de los fenómenos y propuesta metodológicas para la extracción de las UT y rasgos conceptuales identificados  
3.2 No correspondencia entre las variantes de vocabulario y texto

- Otros Ejemplos:

| Variante en Vocabulario normalizado | Variante en Norma texto                    | Fenómeno observado   |
|-------------------------------------|--|--|
| gafa integral                       | gafa con montura integral                  | <b>Variación denominativa:</b> la variante en texto es más explícita que la variante en vocabulario. |
| factor de transmisión               | transmitancia                              | <b>Variación denominativa:</b> Sólo la base léxica de ambas variantes coinciden.                     |
| protector ocular                    | protector ocular;<br>protector de los ojos | <b>Variación denominativa:</b> dos variantes coexisten en los textos. Su base es la misma.           |

**iulaTerm** INSTITUT UNIVERSITARI DE LINGÜÍSTICA APLICADA  
UNIVERSITAT POMPEU FABRA

La definición terminológica en vocabularios normalizados UNE: primera exploración de los datos

3. Presentación de los fenómenos y propuesta metodológicas para la extracción de las UT y rasgos conceptuales identificados  
3.2 No correspondencia entre las variantes de vocabulario y texto

- Pregunta de investigación:

¿Cómo detectar las variantes entre normas de texto y vocabulario normalizado?

**iulaTerm** INSTITUT UNIVERSITARI DE LINGÜÍSTICA APLICADA  
UNIVERSITAT POMPEU FABRA

La definición terminológica en vocabularios normalizados UNE: primera exploración de los datos

3. Presentación de los fenómenos y propuesta metodológicas para la extracción de las UT y rasgos conceptuales identificados  
3.2 No correspondencia entre las variantes de vocabulario y texto

- Estrategias a utilizar:

1. Búsquedas en apartados de la norma susceptibles de contener contextos definitorios o elementos reformuladores (Bach, 2001; Alarcón, 2009): las normas técnicas del ámbito estudiado están estructuradas textualmente sobre el mismo patrón.
2. Búsquedas de los n-gramas más frecuentes en títulos y subtítulos.

**iulaTerm** INSTITUT UNIVERSITARI DE LINGÜÍSTICA APLICADA  
UNIVERSITAT POMPEU FABRA

La definición terminológica en vocabularios normalizados UNE: primera exploración de los datos

3. Presentación de los fenómenos y propuesta metodológicas para la extracción de las UT y rasgos conceptuales identificados  
3.2 No correspondencia entre las variantes de vocabulario y texto

- Metodología:

1. Lista de los patrones morfo-sintácticos de las UT;
2. Marquee estructural del texto;
3. Marquee lingüístico del texto;
4. Extracción de concordancias;
5. Obtención de lista de candidatos a UT que deberían figurar en la norma de vocabulario;
6. Validación por expertos de AENOR.

**iulaTerm** INSTITUT UNIVERSITARI DE LINGÜÍSTICA APLICADA  
UNIVERSITAT POMPEU FABRA

La definición terminológica en vocabularios normalizados UNE: primera exploración de los datos

3. Presentación de los fenómenos y propuesta metodológicas para la extracción de las UT y rasgos conceptuales identificados  
3.3 Contenido conceptual diferente entre UT texto-vocabulario

- Supuesto previo:

Hay información conceptual importante en los textos que no está recogida en los vocabularios.

**iulaTerm** INSTITUT UNIVERSITARI DE LINGÜÍSTICA APLICADA  
UNIVERSITAT POMPEU FABRA

La definición terminológica en vocabularios normalizados UNE: primera exploración de los datos

3. Presentación de los fenómenos y propuesta metodológicas para la extracción de las UT y rasgos conceptuales identificados  
3.3 Contenido conceptual diferente entre UT texto-vocabulario

- Ejemplos:

| UNE-EN 132   | UNE-EN 137  |
|--|---|
| Aire que expulsa el usuario.   | El aire exhalado del usuario pasa <b>sin recirculación</b> a la atmósfera ambiental.  |
| Una válvula controlada por la acción respiratoria de los pulmones, que suministra aire respirable a demanda. | El suministro de aire respirable (de la válvula a demanda) debe ser suficiente para un caudal sinusoidal de 40 x 2,5 l/min a todas las <b>presiones de la botella</b> por encima de 20 bar, y de 25 x 2 l/min a una <b>presión de la botella</b> de 10 bar. |

**iulaTerm** INSTITUT UNIVERSITARI DE LINGÜÍSTICA APLICADA  
UNIVERSITAT POMPEU FABRA

3. Presentación de los fenómenos y propuesta metodológicas para la extracción de las UT y rasgos conceptuales identificados  
3.3 Contenido conceptual diferente entre UT texto-vocabulario

- **Pregunta de investigación:** ¿Cómo identificar los rasgos conceptuales en las normas de texto?
- **Estrategias:**
  1. Búsquedas de reformuladores textuales (Bach, 2001; Alarcón, 2009): la variantes denominativas coexisten textualmente, incluso en textos normativos. Suelen estar introducidas por reformuladores textuales.
  2. Identificación de regularidades semántico-sintácticas de los rasgos conceptuales de los distintas UT en contexto.

3. Presentación de los fenómenos y propuesta metodológicas para la extracción de las UT y rasgos conceptuales identificados  
3.3 Contenido conceptual diferente entre UT texto-vocabulario

- **Metodología:**
  1. Marcaje lingüístico del texto;
  2. Marcaje estructural del texto;

4. Objetivos a corto plazo

- Diseño de la base de datos definitiva;
- Marcaje lingüístico del texto;
- Marcaje estructural del texto;
- Extracción de UT a partir del análisis de reformuladores textuales;

Prueba con una muestra de unidades del corpus de identificación de los rasgos conceptuales a partir del análisis de las unidades del texto con el qualia télico (Pustejovsky, 1991).

[Se observa que una gran parte de las unidades del texto –por su carácter normativo- son sujeto preverbal de verbos de función.]

Gracias.