



NHH

CLARA



**Description and representation
in language resources of Spanish
and English specialized
collocations from Free Trade
Agreements**

**Pedro Patiño García, NHH, Bergen, NO
elda.org, Paris, 26-29.11.2012**



NHH

CLARA



Free Trade Agreements

Pedro Patiño García, NHH, Bergen, NO
elda.org, Paris, 26-29.11.2012

Collocation

One phenomenon, many definitions.

“Lexically and/or pragmatically constrained recurrent co-occurrences of at least two lexical items which are in a direct syntactic relation with each other” (Bartsch, 2004).

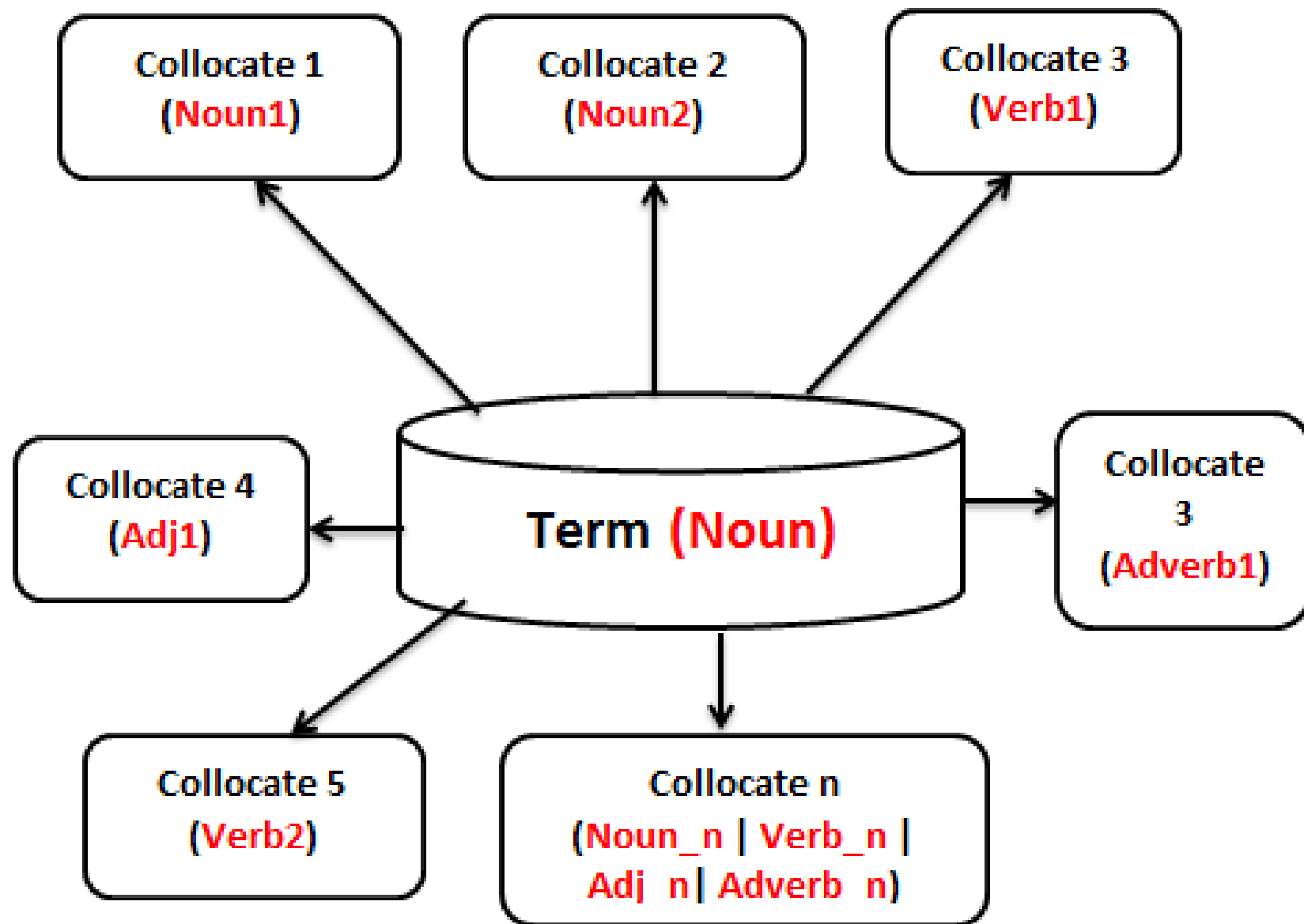
Pedro Patiño García, NHH, Bergen, NO
elda.org, Paris, 26-29.11.2012

Toward a definition of specialised collocation

A specialized collocation is a type of multiword expression composed of at least a **term** that serves as the **node**.

The **collocates** can be **nouns**, **verbs**, **adjectives** or **adverbs** in a direct syntactic relation with the **node**.

These constituents make a **lexical combination** that can be unpredictable and semi-compositional and have an internal and statistical tendency of preference.





Examples of specialised collocations

{reponer | acumular | constituer | arrendar} **capital**;

{increase | apply | favour | pay | raise | adopt | impose | increase | maintain | reduce} **custom duty**

{current | preferred | differential | favorable | free} **tariff treatment**



NHH

CLARA



Specialized collocation:

{**impose** | **extend** | **maintain** | **apply** |
adopt} **a safeguard measure**

{**incrementar** | **adoptar** | **augmentar** |
aplicar | **eliminar** | **reducir**} **un arancel**
aduanero



NHH

CLARA



Method

**Corpus-driven and corpus-based approach.
English-Spanish parallel corpus of FTA texts (1,5 m words per lang.) already aligned with TCA2.**

**Semiautomatic extraction of candidate terms >
candidate specialized collocations**

**Specialized bi- monolingual dictionaries and
termbases like IATE (free trade, economics)**

Reference corpora / methodology

NLP software



NHH

CLARA



Corpus interface

Search with [IMS CWB](#)

30.8.09

Corpus

Book/Chapter

Texts ▾

All ▾

Write book/chapter:

Example: SN.*|SA1|MH10

Word Original (English)		Word Translation (Spanish)	
Whole word ▾	<input type="text" value="support"/>	Whole word ▾	<input type="text"/>
POS	Any ▾	POS	Any ▾
Stop words	<input type="text"/>	Stop words	<input type="text"/>
2nd word			
Whole word ▾	<input type="text"/>	Whole word ▾	<input type="text"/>
POS	Any ▾	POS	Any ▾
3rd word			
Whole word ▾	<input type="text"/>	Whole word ▾	<input type="text"/>
POS	Any ▾	POS	Any ▾

Search

Ignore case

RESET FORM



NHH

CLARA



Corpus interface

Economic Community of 9 February 1970, as amended ; (UE_1992e.s2608)

Estados miembros de la Comunidad Económica Europea el 9 de febrero de 1970, en su versión modificada ; (UE_1992s.s2602)

All claims and liabilities arising from the very short-term financing mechanism and the short-term monetary **support** mechanism, under the Agreements referred to in Article 6.1, shall be settled by the first day of the third stage. (UE_1992e.s2770)

Todos los créditos y deudas resultantes del mecanismo de financiación a muy corto plazo y del mecanismo de ayuda económica a corto plazo estipulados con arreglo a los Acuerdos a que se refiere el artículo 6.1 se satisfarán no más tarde del primer día en que entre en vigor la tercera fase. (UE_1992s.s2763)

1. With a view to achieving the objectives of Article 1, the Community shall **support** and complement the activities of the Member States in the following fields: (UE_1992e.s2914)

1. Para la consecución de los objetivos del artículo 1, la Comunidad apoyará y completará la acción de los Estados miembros en los siguientes ámbitos: (UE_1992s.s2906)

1. The Commission shall have the task of promoting the consultation of management and labour at Community level and shall take any relevant measure to facilitate their dialogue by ensuring balanced **support** for the parties. (UE_1992e.s2934)

1. La Comisión tendrá como cometido fomentar la consulta a los interlocutores sociales a nivel comunitario y adoptar todas las disposiciones necesarias para facilitar su diálogo, velando porque ambas partes reciban un apoyo equilibrado. (UE_1992s.s2926)

- **support** for national criminal investigation and security authorities, in particular in the coordination of investigations and search operations ; (UE_1992e.s3232)

- asistencia a las autoridades nacionales encargadas de la persecución penal y de la seguridad, en particular en materia de coordinación de diligencias e investigaciones ; (UE_1992s.s3224)



NHH

CLARA



IMS Corpus Workbench – CQP

[pos="V.*"] []? "medida" [pos="ADJ"];

FTA-ES> [pos="V.*"] []? "medida" [pos="ADJ"];

20144: guardia textil . La Parte que <adopte la medida arancelaria> la aplicará solamente por el
111943: garantía . 3 El tribunal puede <ordenar una medida provisional> de protección para preservar
127077: Más Favorecida) , deberá ser <tratada como medida disconforme> no sujeta al Artículo 12 . 2
184124: gor de este Acuerdo , e (ii) <incluye cualquier medida subordinada> , adoptada o mantenida bajo
184958: trato nacional y por ello será <listada como medida disconforme> : a un banco extranjero de un
186205: e al sector general al cual se <aplica la medida disconforme> ; (b) Subsector se refiere
186222: l sector específico al cual se <aplica la medida disconforme> ; (c) Obligaciones Afectada
186381: gor de este Tratado , y (ii) <incluye cualquier medida subordinada> , adoptada o mantenida bajo
186478: e al sector general al cual se <aplica la medida disconforme> ; (b) Subsector se refiere
186495: l sector específico al cual se <aplica la medida disconforme> ; (c) Obligaciones Afectada
188681: eserva el derecho de adoptar o <mantener cualquier medida relativa> al suministro de servicios de
193340: gor de este Acuerdo , e (ii) <incluye cualquier medida subordinada> , adoptada o mantenida bajo
381962: á a definir y , en su caso , a <aplicar cualquier medida oportuna> para reforzar la base industr
382881: uir a definir y , en su caso , <aplicar cualquier medida oportuna> para reforzar la base industr
486013: ento de la Unión Europea , que <modifiquen una medida existente> que sea vinculante para dicho
488517: ento de la Unión Europea , que <modifiquen una medida existente> que sea vinculante para dicho
561468: de la Parte importadora podrá <adoptar cualquier medida necesaria> para salvaguardar los interes
586778: iquen que el supuesto subsidio <es una medida gubernamental> de asistencia , directa o ind
594225: desarrollado Parte que decida <aplicar una medida definitiva> en contra de productos origin
621618: sin necesidad de que esa Parte <tome alguna medida legislativa> adicional . No obstante lo di
639130: a sea directa o indirectamente <mediante] medida [s]> equivalentes a la expropiació
639743: inatoria de aplicación general <es una medida equivalente> a la expropiación de un valor
647655: el párrafo 29 . 3 . a) podrá <solicitar una medida provisional> de protección que no incluya
654333: r a las partes contendientes , <[una medida provisional] de protección] [medidas provi
654481:] [44 . 1 . Un Tribunal podrá <recomendar una medida provisional> de protección para preservar



NHH
2 0 1 1 **75**år

CLARA



Need for evaluation

- **Software for term and collocation extraction**
- **Test the suitability of current representation schemes (LMF, TBX, commercial software / private companies) for encoding specialized collocations.**



Encoding of a Spanish phraseological unit using LMF

```
1 <?xml version="1.0" encoding="iso-8859-1"?>
2 <LexicalResource dtdVersion="16">
3 <GlobalInformation>
4   <feat att="label" val="Representación de locuciones"/>
5   <feat att="comment" val="Locución española daños y perjuicios"/>
6   <feat att="languageCoding" val="ISO 639-3"/>
7 </GlobalInformation>
8 <Lexicon>
9   <feat att="language" val="spa"/>
10 <LexicalEntry mwePattern="NConjN">
11   <feat att="partOfSpeech" val="LocNom"/>
12 <Lemma>
13   <feat att="writtenForm" val="daños y perjuicios"/>
14 </Lemma>
15 <ListOfComponents>
16   <Component entry="E1"/>
17   <Component entry="E2"/>
18   <Component entry="E3"/>
19 </ListOfComponents>
20 </LexicalEntry>
21 <LexicalEntry id="E1">
22   <feat att="partOfSpeech" val="nombre"/>
23 <Lemma>
24   <feat att="writtenForm" val="daños"/>
25 </Lemma>
```



NHH

CLARA



References

Bartsch, S. (2004), Structural and functional properties of collocations in English: a corpus study of lexical and pragmatic constraints on lexical co-occurrence, Gunter Narr Verlag, Tübingen.



NHH

CLARA



Thank you!

Your comments are welcome:

pedro.patino@nhh.no