

Developing a Collocation Dictionary Writing System (COLDWS) for an Online Multilingual Collocations Dictionary Platform (PLATCOL)



Adriane ORENHA-OTTAIANO
Maria Eugênia OLÍMPIO
Carlos Roberto VALÊNCIO
João Pedro QUADRADO

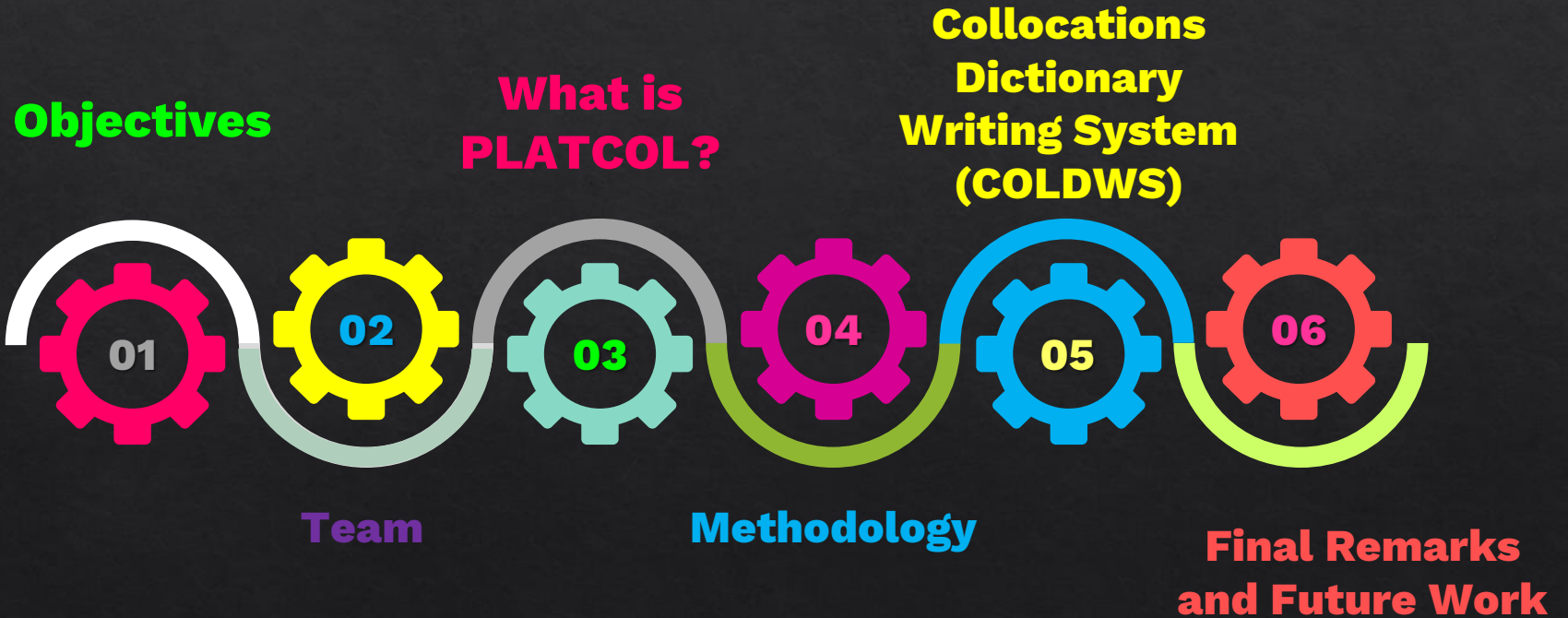


ACKNOWLEDGEMENTS

Tanneke Schoonheim Grant

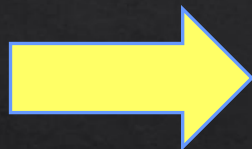


Outline



Project Objectives

It aims at developing a phraseographical methodology and model for an *Online Corpus-Based Multilingual Collocations Dictionary Platform (PLATCOL)*, in five different languages so far.



to be customized for different target audiences according to their needs



Presentation Objectives

We will focus on the development of a COLDWS, duly created to fulfill the specific needs of the collocations dictionaries.

We have knowledge of Lexonomy, a web-based dictionary writing system (Měchura 2017), TshwaneLex (de Schryver 2007), and Fuertes-Olivera & Tarp's (2018), good quality DWSs.

However, we needed to rely on a DWS that would meet the specificities of a Multilingual Collocations Dictionary and also deal with the specific data output generated for our project.



A Phraseographical Methodology and Model for an Online Corpus-Based Multilingual Collocations Dictionary Platform (Process FAPESP 2020/01783-2)



**Profa. Dra. Adriane Orenha
Ottaiano**



**Carlos Roberto Valêncio
João Pedro Quadrado**

**Profa. Dra. Maria
Eugênia Olímpio**

**Profa. Dra. Margarita
Alonso-Ramos**

**Prof. Dr. Marcos
Garcia**



**Profa. Dra. Marie-
Claude L'Homme**



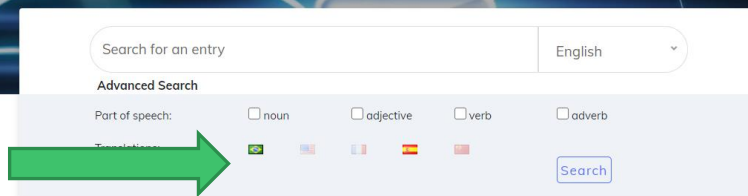
**Prof. Dr. Antonio
Pamies-Bertrán**

**Prof. Dr. José
Manuel Pazos**



Welcome to PLATCOL





A Platform of Multilingual Collocations Dictionaries



Search for an entry English

Advanced Search

Part of speech: noun adjective verb adverb

Translations:    

PLATCOL

PLATCOL is an Online Platform of Multilingual Collocations Dictionaries. It aims to promote learning and translation of collocations more effectively, so that the dictionaries' users can develop proficiency and fluency as well as achieve native-like naturalness in many different languages.



About the PLATCOL platform

PLATCOL is an Online Platform of Multilingual Collocations Dictionaries. It aims to promote learning and translation of collocations more effectively, so that the dictionaries' users can develop proficiency and fluency as well as achieve native-like naturalness in many different languages. So far PLATCOL comprises of five languages: English, Portuguese, French, Spanish and Chinese. However, more languages will be coming soon.

About us

[Team Members](#)
[Others platforms](#)
[Contact](#)
[Publications](#)

Quick Links

[What are collocations?](#)
[Register](#)
[Login](#)
[References](#)



São Paulo Research Foundation



São Paulo State University

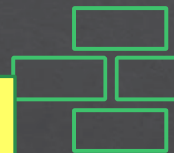
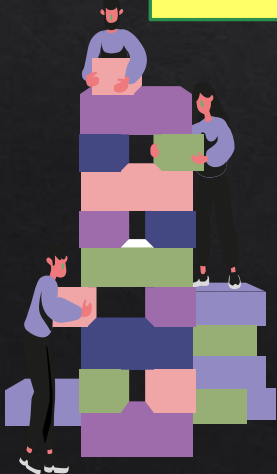
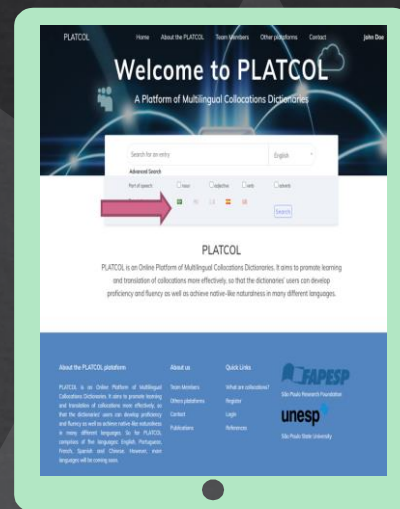
<http://www.institucional.grupogbd.com/dicionario/sobre?locale=en>

What is PLATCOL?

❖ A Platform that aims to offer a new way to search collocations, with a new dictionary format

❖ A Platform that seeks to meet its users' needs - higher customization of the structure of the dictionaries

❖ Production dictionaries that can provide help to their users and contribute to the development of their collocational knowledge and competence



Our proposal is based on the Functional Theory of Lexicography (Bergenholtz & Tarp 2003, Bothma & Tarp 2012, Fuertes-Olivera & Tarp 2014, Tarp 2013, 2014 and 2015).

Thus, both the procedures chosen for the selection, organization and presentation of lexicographical data, as well as the determination of the content, form and access routes are adapted and subordinated to the preferences of the users.



USERS



**Pre-service
Teachers**



In-service Teachers



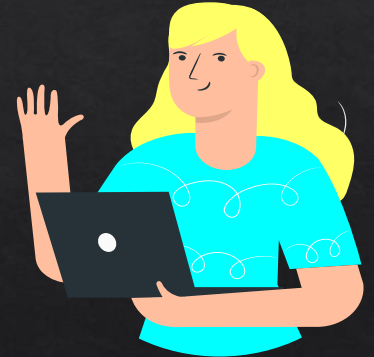
**Learner and Professional
Translators**



Language Learners



**Researchers and
Lexicographers**



Material Developers

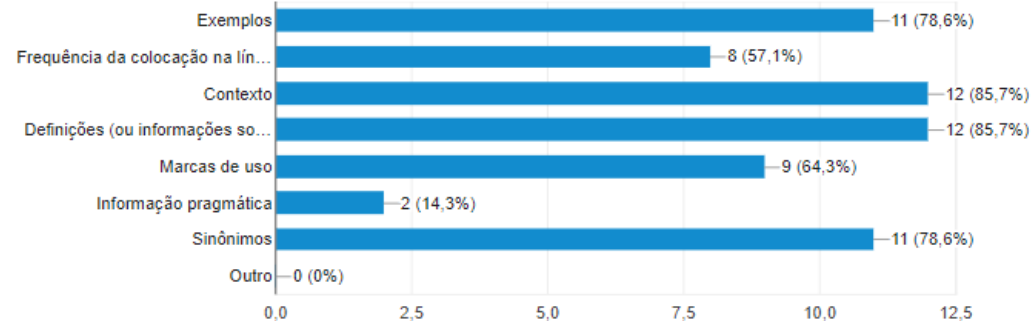
Survey on users' needs



Que tipo de informação você esperaria de um dicionário de colocação de língua geral monolíngue?

[Copiar](#)

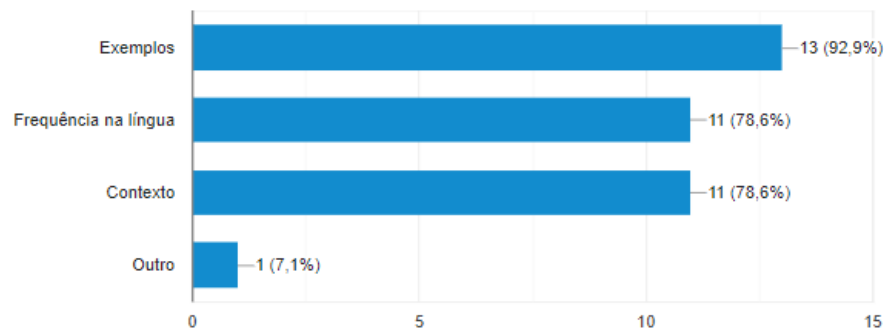
14 respostas



Que tipo de informação você esperaria de um dicionário de colocações de língua geral Bilingue ou Multilíngue?

[Copiar](#)

14 respostas



METHODOLOGY

The methodology is based on the automatic approach described in Garcia *et al.* (2019a) and a manual review and validation of the extracted data made by lexicographers.



THE CORPUS

A large corpus for each of the five languages was automatically compiled, using different source data – and parsed with UDPipe (Straka & Straková, 2017):

Language	Sources	Size (tokens)
Portuguese	Jornal do Brasil, Wikipedia/Wikibooks, Paracrawl, CHAVE (Santos & Rocha, 2004), CBras, BrWaC (Wagner Filho et al., 2018)	4B
Spanish	EuroParl (Kohen, 2005), Literature (short stories/romances) (Garcia et al., 2019a), Wikipedia/Wikibooks	1,2B
English	EuroParl, Wikipedia/Wikibooks	1.6B
French	FrWaC (Baroni et al., 2009), Wikipedia/Wikibooks	2.5 B
Chinese	Wikipedia, Wikibooks, and literary texts	600M

METHODOLOGY

**Automatic
retrieval of
bases**

Focus on collocation types with three
morphosyntactic classes of bases:

- ☐ nouns
- ☐ verbs
- ☐ adjectives



AUTOMATICALLY EXTRACTED BASES

	French	Portuguese	English	Spanish	Chinese
Nouns	9,754	10,307	10,545	8,853	1,858
Verbs	4,895	5,573	5,502	3,549	2,206

VALIDATED CANDIDATES

	French	Portuguese	English	Spanish	Chinese
Nouns	8,361	8,690	8,713	6,576	103
Verbs	2,902	3,817	3,982	2,459	100

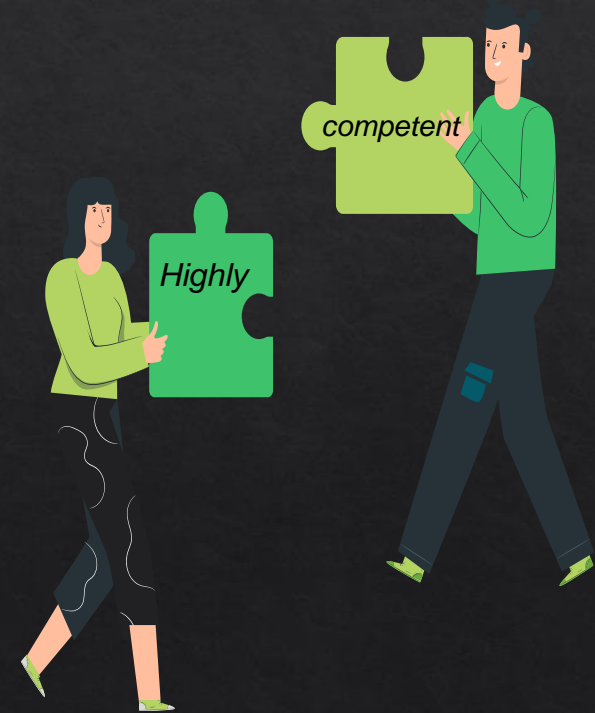
**VALIDATED
BASES**



METHODOLOGY

Automatic Extraction of Collocations

- We extracted dependency triples with the selected bases, and organize them as candidates with the following structure: *relation (base, collocate)*.



Registered Collocaciones in the *COLDWS*



Chinese	5,786
Spanish	71,131
French	60,447
English	52,826
Portuguese	62,968
Total	253,158

Identification of example sentences

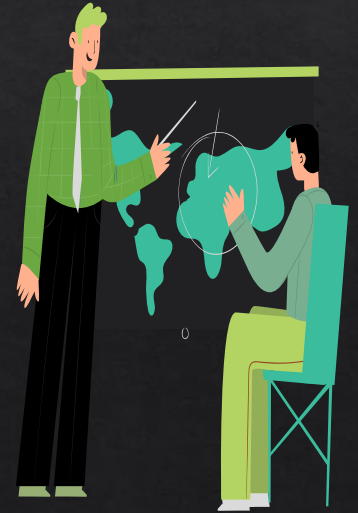
❖ We use GDEX-inspired heuristics (Kilgariff *et al.*, 2008) to automatically assign a score to each sentence in the corpus (containing candidate collocations).

❖ Following Kosem *et al* (2019), we discard sentences with less than 6 tokens and incrementally penalize those with more than 30 tokens.



METHODOLOGY

**Automatic
Translation of
Collocations**
(Garcia *et al.* 2019c)





Electronic lexicography in the 21st century: post-editing lexicography

Proceedings of the eLex 2021 conference

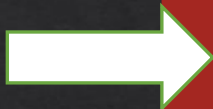
edited by

Iztok Kosem
Michal Culr
Miloš Jakubiček
Jelena Kallas
Simon Krek
Carole Tiberius

virtual, 5–7 July 2021

elex.link/elex2021

➤ For more information



Corpus-based Methodology for an Online Multilingual Collocations Dictionary: First Steps

Adriane Orenha-Ottaiano¹, Marcos Garcia², Maria Eugênia
Olimpio de Oliveira Silva³, Marie-Claude L'Homme⁴, Margarita
Alonso Ramos⁵, Carlos Roberto Valêncio⁶, William Tenório⁷

¹ São Paulo State University (UNESP), Brazil

² Universidade de Santiago de Compostela, Galiza, Spain

³ University of Alcalá, Spain

⁴ OLST, Université de Montréal, Québec, Canada

⁵ Universidade da Coruña, Spain

⁶ São Paulo State University (UNESP), Brazil

⁷ São Paulo State University (UNESP), Brazil

E-mail: adriane.ottaiano@unesp.br, marcos.garcia.gonzalez@usc.gal, eugenia.olimpio@uah.es,
mc.lhomme@umontreal.ca.ca, margarita.alonso@ude.es, carlos.valencio@unesp.br,
williamtenoriotenorio@gmail.com

Abstract

This paper describes the first steps of a corpus-based methodology for the development of an online Platform for Multilingual Collocations Dictionaries (PLATCOL). The platform is aimed to be customized for different target audiences according to their needs. It covers various syntactic structures of collocations that fit into the following taxonomy: verbal, adjectival, nominal, and adverbial. Part of its design, layout and methodological procedures are based on the *Bilingual Online Collocations Dictionary Platform* (Orenha-Ottaiano, 2017). The methodology also relies on the combination of automatic methods to extract candidate collocations (Garcia et al., 2019a) with careful post-editing performed by lexicographers. The automatic approaches take advantage of NLP tools to annotate large corpora with lemmas, PoS-tags and dependency relations in five languages (English, French, Portuguese, Spanish and Chinese). Using these data, we apply statistical measures (Evert et al., 2017; Garcia et al., 2019b) and distributional semantics strategies to select the candidates (Garcia et al., 2019c) and retrieve corpus-based examples (Kilgarriff et al., 2008). We also rely on automatic definition extraction (Bond & Foster, 2013) so that collocations can be more effectively organized according to their specific senses.

Keywords: collocations; collocations dictionary; online platform; automatic extraction; lexicography

1. Introduction

In the past two decades, collocations have been high on the agenda of foreign language teaching and learning (Nesselhauf, 2005; Alonso-Ramos, 2008, 2019; Laufer, 2011; Orenha-Ottaiano, 2021; Torner & Bernal, 2017, among others). Despite this fact, when it comes to the translation of collocations, the number of studies that can contribute to better comprehension of the difficulties regarding the complexity of translation of such combinations is not as significant (Kenny, 2001; Bernardini, 2007; Gregorio-Godeo & Molina, 2011; Orenha-Ottaiano, 2009, 2012, forthcoming).

Signature

noun

1. your name written in your own handwriting as a way of identification for a document, check etc., making it difficult to be copied

verb + SIGNATURE collocations

collect signature

Arizona petition circulators had to change their operations to collect **signature** and meet the July deadline for 2020 ballot initiatives. [See more](#)

forge signature

Leading loyalist campaigner Jamie Bryson appears to have **forged a signature** used to verify a set of annual accounts for a community group where he works, according to one of the UK's leading hand-writing experts. [See more](#)

gather signature

More also said officials should allow activists to **gather signatures** online. But that's currently not allowed in any state, said Amanda Zoch, who works on elections at the National Conference of State Legislatures. [See more](#)

validate signature

In short, she had **validated her signature** on two letters that were specifically stated as confidential in an agreement we had both signed. [See more](#)

adjective + SIGNATURE collocations

digital signature

In California, activists working on measures to legalize psilocybin mushrooms and amend the state's legal marijuana program have pushed for **digital signature** collection options. [See more](#)

electronic signature

Today, there are some states that accept **electronic signatures** and remote online notarization for real estate transactions and the recording or filing of documents. [See more](#)

valid signature

Envelopes with a **valid signature** are cut open using an extraction machine, then the ballots themselves are scanned and counted. Izary says the paper trail prevents hacking and allows for a verifiable archive that can later be double-checked. [See more](#)

Other entries

Sign noun
Sign verb

Advanced options

- Show collocation taxonomy
- Show collocation frequency
- Show collocation statistics
- Show translations

About the PLATCOL platform

PLATCOL is an Online Platform of Multilingual Collocations Dictionaries. It aims to promote learning and translation of collocations more effectively, so that the dictionary users can develop proficiency and fluency as well as achieve native-like naturalness in many different languages. So far PLATCOL comprises of five languages: English, Portuguese, French, Spanish and Chinese. However, more languages will be coming soon.

About us

Team Members
Others platforms
Contact
Publications

Quick Links

What are collocations?
Register
Login
References



São Paulo Research Foundation



São Paulo State University



Sign (Freq. 2,666,850)

verb

1. to write your signature on something to show that you wrote it, created it, or agreed with it

adv. + SIGN collocations

digitally sign (Freq. 4,652) (TS 65.61 - MI 18.25 - LogDice 10.76)

Alice selects the file to be digitally signed or clicks on 'sign' in her email application. [See more](#)

duly sign (Freq. 2,356) (TS 48.08 - MI 17.86 - LogDice 9.44)

In 2002, they took out a mortgage with The Woolwich (part of Barclays), and were sent a deed of mortgage which they duly signed. [See more](#)

jointly sign (Freq. 738) (TS 26.93 - MI 17.00 - LogDice 6.93)

Partnership shall be registered by way of the partners jointly signing the partnership document in the presence of the authority referred to in section 4(1) and that authority verifying the same with a countersignature. [See more](#)

Other entries

Sign noun
Signature noun

Advanced options

- Show collocation taxonomy
- Show collocation frequency
- Show collocation statistics
- T-Score
- Mutual Information
- LogDice
- Show translations

✓ Advanced options will be available if a user opts to sign in, according to a users' profile.

✓ A new dictionary structure will be available as users can choose from items in a Menu



About the PLATCOL platform

PLATCOL is an Online Platform of Multilingual Collocations Dictionaries. It aims to promote learning and translation of collocations more effectively, so that the dictionaries' users can develop proficiency and fluency as well as achieve native-like naturalness in many different languages. So far PLATCOL comprises of five languages: English, Portuguese, French, Spanish and Chinese. However, more languages will be coming soon.

About us

[Team Members](#)
[Others platforms](#)
[Contact](#)
[Publications](#)

Quick Links

[What are collocations?](#)
[Register](#)
[Login](#)
[References](#)



São Paulo Research Foundation



São Paulo State University

Collocations Dictionary Writing System (COLDWS) - Purposes



PLATCOL

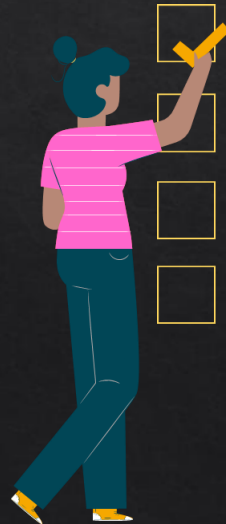
Enter your email and password

Enter

[Recover password](#)

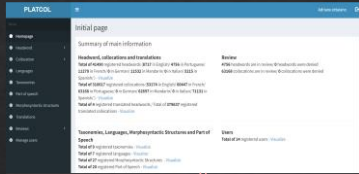
unesp
UNIVERSIDADE ESTADUAL PAULISTA
"JÚLIO DE MESQUITA FILHO"

gbd
Grupo de Banco de Dados
unesp



- ✓ Better enable or enhance the post-editing of all the automatically retrieved data
- ✓ Developed to specifically compile and produce collocation dictionaries, so that all automatically extracted data are automatically inserted into this COLDWS, post-edited by the lexicographers, as well as be exported to an end-user platform.

Collocations Dictionary Writing System (COLDWS)

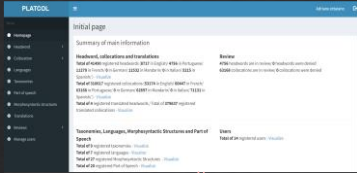


The software has a lot of functionalities – register, edit, and post-edit all dictionary information, such as:

- senses
- collocations
- morphosyntactic structures
- taxonomy of the collocations
- Translations
- statistical measures etc.



Collocations Dictionary Writing System (COLDWS)



In what concerns computational development, the COLDWS was built using languages from current and widespread programming such as Java (with Model View Controller [MVC] architecture), HTML and Javascript (jQuery library)

For the storage, consultation and deletion of entries, the relational data model with PostgreSQL software was used in conjunction with SQL language

In addition, concepts of User Experience - UX were applied to provide a good experience for the lexicographers.



COLDWS will also be of valuable help when it comes to optimizing translation of entries or collocations

COLDWS – Translations

For example, the collocations “develop a plan”, in English, and “desenvolver um plano”, in Portuguese, are equivalents. Similarly, the collocations “desenvolver um plano”, in Portuguese, and “desarrollar un plan”, in Spanish, also have a translation relationship.

This way, even if it has not been previously identified in the automatic extraction process, the relationship between the collocations “develop a plan” and “desarrollar un plan” will be automatically inferred.



It shows lexicographers
the possibilities of
translations

They analyze the
reliability and quality of
the translation found
(from 0 -1).

Collocation	Translation	Trust value		
fazer diferença	faire différence (French)	0.98	Approve	Disapprove
fazer diferença	hacer diferencia (Spanish)	0.90	Approve	Disapprove
fazer diferença	make difference (English)	0.93	Approve	Disapprove
ter condição	devoir condition (French)	0.87	Approve	Disapprove
ter condição	hacer condición (Spanish)	0.89	Approve	Disapprove
ter condição	have condition (English)	0.94	Approve	Disapprove
ter opção	pouvoir option (French)	0.97	Approve	Disapprove
ter opção	tener opción (Spanish)	0.96	Approve	Disapprove
ter opção	give option (English)	0.72	Approve	Disapprove

Collocations Dictionary Writing System (COLDWS)

The screenshot shows the FLATCOL website interface. On the left is a dark sidebar menu with the following items: Home page, Header, Collocations, Languages, Taxonomies, Part of speech, Morphosyntactic structures, Translations, Users, and Manage users. The main content area is titled 'Initial page' and contains a 'Summary of main information' section. This section is divided into three columns of statistics:

- Headword, collocations and translations:**
 - Total of 40490 registered headwords: 3117 in English, 4794 in Portuguese, 21179 in French, 0 in German, 11503 in Mandarin, 0 in Italian, 1115 in Spanish. [Visualize](#)
 - Total of 518617 registered collocations: 53174 in English, 60447 in French, 63158 in Portuguese, 0 in German, 42887 in Mandarin, 0 in Italian, 71131 in Spanish. [Visualize](#)
 - Total of 0 registered translated headwords, Total of 278627 registered translated collocations. [Visualize](#)
- Taxonomies, Languages, Morphosyntactic Structures and Part of Speech:**
 - Total of 9 registered taxonomies. [Visualize](#)
 - Total of 7 registered languages. [Visualize](#)
 - Total of 27 registered Morphosyntactic Structures. [Visualize](#)
 - Total of 29 registered Part of Speech. [Visualize](#)
- Review:**
 - 4794 headwords are in review, 0 headwords were denied.
 - 62068 collocations are in review, 0 collocations were denied.
- Users:**
 - Total of 34 registered users. [Visualize](#)



Menu

- Homepage
- Headword
- Register Headword
- **Headwords**
- Collocation
- Languages
- Taxonomies
- Part of speech
- Morphosyntactic structures
- Translations
- Reviews
- Manage users

Headwords

Register new headword (F10)

List of headwords

Mostrar 10 registros

Buscar:

#	Cod.	Headword	Part of speech	Status	Number of meanings	Number of collocations
...	9500	abaissement	nom	●	1	0
...	9501	abandon	nom	●	1	0
...	9502	abattage	nom	●	1	0
...	9503	abattement	nom	●	1	0
...	9504	abattoir	nom	●	1	0
...	9505	abbatiale	nom	●	1	0

Menu

- Homepage
- **Headword** ▾
- Register Headword
- Headwords
- Collocation <
- Languages
- Taxonomies
- Part of speech
- Morphosyntactic structures
- Translations

Register Headword

Headword Information

Senses

Corresponding headwords and Comments

Headword language

Headword

Frequency

Part of Speech ⓘ

 Headword Part of Speech - ATTENTION: : Please first type the headword language

Page 1 out of 3

[Advance page](#)

Register Headword

Headword Information

Senses

Corresponding headwords and Comments

 Does it have Only one sense?

Senses

Page 2 out of 3

Register Headword

Headword Information

Senses

Corresponding headwords and Comments

Corresponding Headwords

Comments

 Comments about the Headword

Page 3 out of 3

[Go back to previous page](#)

Menu

- Homepage
- Headword <
- Collocation >
- Register collocation
- Collocations
- Languages
- Taxonomies
- Part of speech
- Morphosyntactic structures
- Translations
- Reviews <
- Manage users

Register collocation

Collocation information

Examples

Corresponding collocations, Comments and Observations

Headword meaning related to the collocation

Only one meaning

Collocation

Collocation

Collocate

Collocate

Base

Justice

Part of speech

Collocate part of speech

Frequency

Collocation frequency

Mutual Information

4.41289857012632

Delta pi

0.0196665241556221

SL

5424.36756414534

Log Dice

157.864151728346

LL

34.2039372944186

ZS

2804.43259119987

Synonymous collocations

Collocations notes

Usage notes

Usage notes

Lexicographical remarks


Lexicographical remarks

Edit collocation

Collocation information

Examples

Corresponding collocations, Translation notes and Observations

Under revision 

Example 1

Bold *Italic* Underline



Por certo que esse reconhecimento não está restrito a as **peessoas jurídicas** de direito privado .

Example source

<https://www.google.com/>

Example 2

Bold *Italic* Underline



Por outro lado , por mais redundante que possa soar , os entes públicos , enquanto pessoas jurídicas , são pessoas .




Collocations

Edit collocation

Collocation information

Examples

Corresponding collocations, Translation notes and Observations

Under revision 

Corresponding collocations

Language

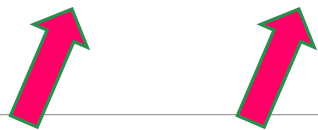
Spanish

Corresponding collocation

persona jurídica



- Free combination A single lexical unit



Add corresponding collocation

Observations

Observations about the collocation.

Translation Notes

Notes about the translations.

LEGAL PERSON

- Homepage
- Headword <
- Collocation ▾
 - Register collocation
 - Collocations
- Languages
- Taxonomies
- Part of speech
- Morphosyntactic structures
- Translations
- Reviews <
- Manage users

Collocations

Register new collocation (F10)

List of collocations

Mostrar 10 registros

Buscar:

#	↑↓	Cod.	↕	Collocation	↑↓	Part of speech	↑↓	Status	↑↓	Number of corresponding collocations	↑↓	Related to the headword	↑↓
...		20		amar loucamente		verbo		●		0		amar	
...		2000		peessoa jurídica		substantivo		●		1		peessoa	
...		2002		fazer parte		substantivo		●		2		parte	
...		2003		último ano		substantivo		●		3		ano	
...		2004		grande parte		substantivo		●		2		parte	
...		2006		ter direito		substantivo		●		4		direito	
...		2007		entrar contato		substantivo		●		3		contato	
...		2013		primeiro ano		substantivo		●		1		ano	
...		2014		rede social		substantivo		●		1		rede	
...		2015		chamar atenção		substantivo		●		4		atenção	

PLATCOL Giseli Cecilio

Menu

- Homepage
- Headword
- Collocation
- Register collocation
- Collocations
- Taxonomies
- Part of speech
- Morphosyntactic structures
- Translations
- Reviews

Edit collocation

Collocation information | Examples | Corresponding collocations, Translation notes and Observations

Collocation information

Headword sense related to the collocation

Only one sense

Collocation

compose song

Collocate

compose

Base

song

Part of speech

noun

Frequency

2505

Mutual Information

4.65972690898978

Taxonomy

No taxonomy

Complement

Complement

Morphosyntactic Structure

No morphosyntactic structure

T-Score

0.0166217720882193

Log Dice

241.675047605675

Under revision ●

With respect to the validation process, the software will allow lexicographers to choose from three phases (traffic lights phases), indicating to users the status of the entries or collocations.



Phase 1 – data automatically inserted into the COLDWS (not revised yet): a red icon will be displayed, even to users, when exported to the end-user platform



Phase 2 – represents data revised by one member of the team (reviewer 1), but which may still need a second evaluation and/or some adjustments: an orange icon will be shown.



Phase 3 – data is checked by a second reviewer (reviewer 2) and now considered to be suitable

Menu

- Homepage
- Headword <
- Collocation ▾
- Register collocation
- Collocations
- Taxonomies
- Part of speech
- Morphosyntactic structures
- Translations
- Reviews <

song

Headword language: English
Frequency: 675400

Headword part of speech: noun

Edit collocation

Visualize collocation on
the website

Delete collocation

unreleased song

| Under review



Language: English
Collocate: unreleased
Taxonomy: No taxonomy
Part of speech: noun
Morphosyntactic structure: No morphosyntactic structure
Delta Pi: 0.0206258495852637
LL: 38.0714709616693

Complement:
Frequency: 1487
T-Score: 0.0107903549004168
Mutual Information: 6.29774623565067
Log Dice: 337.679672476587
SL: 10046.1016136351
ZS: 5545.32343757153

Related to sense
Only one meaning

Usage notes
No usage notes

Menu

- Homepage
- Headword <
- Collocation >
- Register collocation
- Collocations
- Taxonomies
- Part of speech
- Morphosyntactic structures
- Translations
- Reviews <

song

Headword language: English
Frequency: 675400

Headword part of speech: noun



Edit collocation



Visualize collocation on
the website



Delete collocation

famous song

Approved ●

Language: English
Collocate: famous
Taxonomy: Adjectival
Part of speech: noun
Morphosyntactic structure: Adjective + Noun
Delta Pi: 0.0150001368625626
LL: 35.6013098710351

Complement: No complement
Frequency: 1644
T-Score: 0.0105973930282427
Mutual Information: 3.0355384424695
Log Dice: 101.943617184775
SL: 4031.19767835903
ZS: 2083.1210296154

Related to sense
Only one meaning

Usage notes
No usage notes

PLATCOL

Menu

- Homepage
- Headword
- Collocation
- Languages
- Taxonomies
- Part of speech
- Morphosyntactic structures
- Translations
- Reviews
- Manage users

Languages

List of Languages

Mostrar 10 registros

Cod.	Language
1	Portuguese
2	English
3	Spanish
4	French
5	Mandarin
9	Italian
10	German

PLATCOL

Menu

- Homepage
- Headword
- Collocation
- Languages
- Taxonomies
- Part of speech
- Morphosyntactic structures
- Translations
- Reviews
- Manage users

English Adriane ottaiano

Morphosyntactic structures

Register new morphosyntactic structure (F10)

List of Morphosyntactic structures

Mostrar 10 registros

Buscar:

Cod.	Morphosyntactic structure	Language of morphosyntactic structures
5	Noun + Verb	English
6	Adjective + Noun	English
7	Adverb + Adjective	English
8	Adverb + Verb	English
16	Verb + Noun	English
22	Verb + Adverb	English
31	No morphosyntactic structure	English

English Adriane ottaiano

Taxonomy

Register new taxonomy (F10)

List of taxonomies

Mostrar 10 registros

Buscar:

Cod.	Taxonomy name	Taxonomy language
1	No taxonomy	English
2	Verbal	English
3	Nominal	English
4	Adjectival	English
5	Adverbial	English

Mostrando 1 até 5 de 5 registros

Anterior 1 Próxima

Menu

- Homepage
- Headword <
- Collocation <
- Languages
- Taxonomies
- Part of speech
- Morphosyntactic structures
- **Translations**
- Reviews <
- Manage users

Translations

List of translations

Mostrar registrosBuscar:

#	⇅	Cod.	📄	Name	⇅	Translation	⇅	Translation language	⇅	Type	⇅	Translation status	⇅
...		9003		sign		assinar		Portuguese		Headword		●	
...		67001		organic chemistry		有机化学		Mandarin		Collocation		●	
...		67001		organic chemistry		orgánico química		Spanish		Collocation		●	
...		67001		organic chemistry		chimie organique		French		Collocation		●	
...		67003		organic molecule		orgánico molécula		Spanish		Collocation		●	
...		67004		page talk		página discusión		Spanish		Collocation		●	
...		67004		page talk		comentário blog		Portuguese		Collocation		●	
...		67007		time interval		intervalo tiempo		Spanish		Collocation		●	
...		67007		time interval		intervalo tempo		Portuguese		Collocation		●	
...		67007		time interval		temps intervalle		French		Collocation		●	

- Homepage
- Headword
- Collocation
- Languages
- Taxonomies
- Part of speech
- Morphosyntactic structures
- Translations
- Reviews
- **Manage users**







Users

Register new user (F10)

List of users

Mostrar 10 registros

















Buscar:

Cod.	Name	Email	Status of User	Type of user	
1	Adriane ottaiano	adriane.ottaiano@unesp.br	Active user	Main Administrator	 
2	GBD Infra	gbd.infra@gmail.comGBD	Active user	Main Administrator	 
3	João Pedro				 

List of users

Mostrar 10 registros

Buscar:

Cod.	Name	Email	Status of User	Type of user	
16	Lais Real	lais.real@unesp.br	Active user	Student	 
17	Emanoel Alves	emanoel.alves@unesp.br	Active user	Student	 
18	Ariane Caldas	ariane.caldas@unesp.br	Active user	Student	 
19	Leila Felipini	leila.felipini@unisagrado.edu.br	Active user	Student	 
20	Daniel Meneses de Souza	dm.souza@unesp.br	Active user	Main Administrator	 
22	Daniel Meneses	danielmsms1995@gmail.com	Active user	Student	 
23	Luiz Carlos Alves Junior	luiz.alves.jr.2019@gmail.com	Active user	Student	 
24	Beatriz Martinez Rossi	beatriz.rossi@unesp.br	Active user	Student	 

Mostrando 11 até 18 de 18 registros

















Anterior 1 2 Próxima

- Homepage
- Headword <
- Collocation <
- Languages
- Taxonomies
- Part of speech
- Morphosyntactic structures
- Translations
- **Reviews** ▾
 - My Reviews
 - **Manage headwords**
 - Manage collocations
 - Manage headwords translations
 - Manage collocations translations
- Manage users

Manage headwords under review

List of headwords under review

Mostrar registros

Cod.	Headword	Reviewer email 1	Reviewer email 2	Headword status	Frequency	
3079	abbreviation	Undefined	Undefined	●	9427	 
4932	Abc	Undefined	Undefined	●	4288	 
6887	abdication	Undefined	Undefined	●	2464	 
4317	abdomen	Undefined	Undefined	●	5371	 
4831	abduction	Undefined	Undefined	●	4442	 
7180	aberration	Undefined	Undefined	●	2282	 
227	ability	Undefined	Undefined	●	148558	 
4259	abnormality	Undefined	Undefined	●	5504	 

There is also a functionality with which reviewers can post-edit entries, collocations and translations

PLATCOL Giseli Cecílio

Menu

- Homepage
- Headword <
- Collocation <
- Taxonomies
- Part of speech
- Morphosyntactic structures
- Translations
- Reviews ▾
- My Reviews

99721	film song	Collocation	Awaiting review ⌚	Review
100739	recording song	Collocation	Awaiting review ⌚	Review
100775	play song	Collocation	Awaiting review ⌚	Review
101692	only song	Collocation	Awaiting review ⌚	Review
101542	receive song	Collocation	Awaiting review ⌚	Review
101706	dance song	Collocation	Awaiting review ⌚	Review
101946	feature song	Collocation	Awaiting review ⌚	Review
102014	successful song	Collocation	Awaiting review ⌚	Review
102126	describe song	Collocation	Awaiting review ⌚	Review

Mostrando 111 até 120 de 3,312 registros 1 row selected

Anterior 1 ... 11 **12** 13 ... 332 Próxima

Reviewer's
screen

Menu

- Homepage
- Headword <
- Collocation <
- Taxonomies
- Part of speech
- Morphosyntactic structures
- Translations
- Reviews ▾
- My Reviews

[Approve collocation](#)[Disapprove collocation](#)[Edit collocation](#)

feature song

Language: English

Collocate: feature

Taxonomy: No taxonomy

Part of speech: noun

Morphosyntactic structure: No morphosyntactic structure

Delta Pi: 0.0227624802869384

LL: 31.691804915703

Frequency: 1631

T-Score: 0.0194172172671808

Mutual Information: 2.21577763631343

Log Dice: 68.3053529869156

Complement:

SL: 2450.18767693624

ZS 1309.20836728811

Related to meaning

Only one meaning

Usage notes

No usage notes

Menu

● Homepage

● Headword <

● Collocation <

● Taxonomies

● Part of speech

● Morphosyntactic structures

● Translations

● Reviews ▾

○ My Reviews

Approve collocation

Disapprove collocation



Edit collocation

feature song

**Approve collocation**

Are you sure you want to approve the collocation feature song? You can only review it once.

No

Yes

SL: 2450.18767693624

ZS 1309.20836728811

famous song

Language: English



Collocation approved

The collocation was successfully approved.

OK

Complement: No complement

SL: 4031.19767835903

ZS 2083.1210296154

Related to meaning

Only one meaning

Usage notes

No usage notes

Menu

- Homepage
- Headword <
- Collocation <
- Taxonomies
- Part of speech
- Morphosyntactic structures
- Translations
- Reviews ▾
- My Reviews

[Approve collocation](#)[Disapprove collocation](#)[Edit collocation](#)

feature song

Language: English**Collocate:** feature**Taxonomy:** No taxonomy**Part of speech:** noun**Morphosyntactic structure:** No morphosyntactic structure**Delta Pi:** 0.0227624802869384**LL:** 31.691804915703**Frequency:** 1631**T-Score:** 0.0194172172671808**Mutual Information:**

2.21577763631343

Log Dice: 68.3053529869156**Complement:****SL:** 2450.18767693624**ZS** 1309.20836728811**Related to meaning**

Only one meaning

Usage notes

No usage notes

Useful notes

No useful notes

Menu

● Homepage

● Headword

● Collocation

● Taxonomies

● Part of speech

● Morphosyntax

● Translations

● Reviews

○ My Reviews

Disapprove collocation ✕**Enter the reason for disapproving the collocation**

Disapproval reason

Close

Disapprove collocation

We are already preparing some options so that the reviewer can just click on one of them, instead of having to write whenever he or she has to disapprove it.

Frequency: 1631

T-Score: 0.0194172172671808

Mutual Information: 2.21577763631343

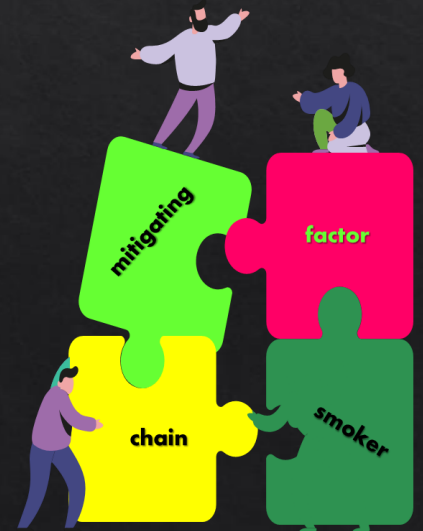
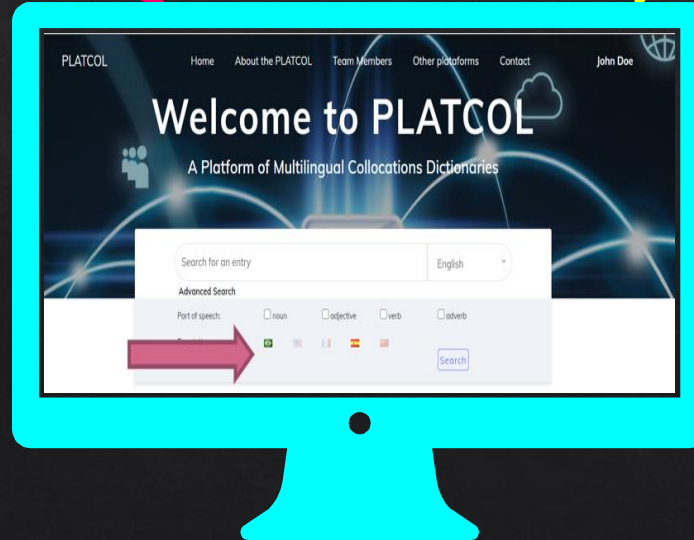
Log Dice: 68.3053529869156

Complement:

CONCLUDING REMARKS

We aim to objectively establish an appropriate methodology and phraseographic model for PLATCOL.

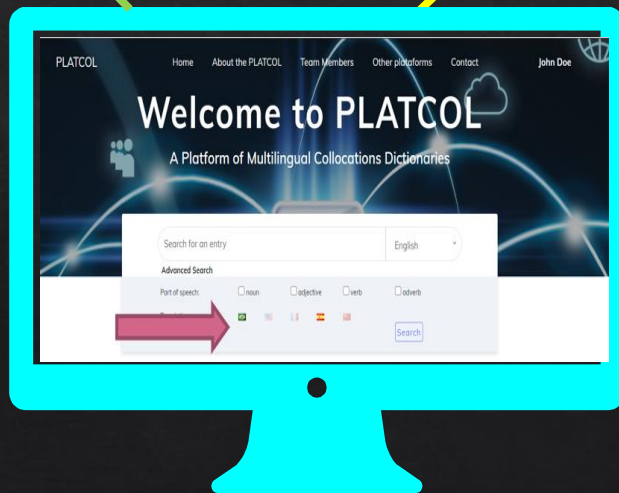
Likewise, we intend to contribute to the development of the collocational competence of its users.



FUTURE WORK

Need for carrying out some minor improvements in the COLDWS.

Some macro and microstructure decisions may still be reshaped, with a view to best adjust the dictionaries to the new languages as well as to users' different lexicographical needs – will be further discussed in future work.



Finish developing the end-user platform in order to launch PLATCOL until the end of 2022!

COLDWS – Translations into other languages

Once translation pairs between collocations are identified and registered in the system, making up a multilingual database, it becomes possible to identify and automatically suggest new translations among other languages (Mausam et al., 2010).



COLDWS – Translations into other languages

This process occurs through an inference-based algorithm (e.g. Mausam et al., 2010), built from an inference hypothesis related to the composition of multiple translation dictionaries:

If word A translates into word B which in turn translates into word C, what is the probability of C being a translation of A?

