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
What is PLATCOL?

Collocations Dictionary Writing System (COLDWS)

Objectives

Team Methodology

Final Remarks and Future Work
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
We will focus on the development of a COLDWS, duly created to fulfill the specific needs of the collocations dictionaries.

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.

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.
A Phraseographical Methodology and Model for an Online Corpus-Based Multilingual Collocations Dictionary Platform (Process FAPESP 2020/01783-2)
Welcome to PLATCOL
A Platform of Multilingual Collocations Dictionaries

Search for an entry

Advanced Search
Part of speech:
- noun
- adjective
- verb
- adverb

Dictionaries:

Search

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. So far, PLATCOL comprises of five languages: English, Portuguese, French, Spanish and Chinese. However, more languages will be coming soon.

About the PLATCOL platform
About us
- Team Members
- Others platforms
- Contact
- Publications

Quick Links
- What are collocations?
- Register
- Login
- References

São Paulo Research Foundation
São Paulo State University

Database Research Group - GBD / São Paulo State University, Brazil. 2021

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

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.
Material Developers

Developers

Language

Learners

Pre-service Teachers

In-service Teachers

Learner and Professional Translators

Researchers andLexicographers

Material Developers
Survey on users’ needs
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.
A large corpus for each of the five languages was automatically compiled, using different source data – and parsed with UDpipe (Straka & Straková, 2017):

<table>
<thead>
<tr>
<th>Language</th>
<th>Sources</th>
<th>Size (tokens)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Portuguese</td>
<td>Jornal do Brasil, Wikipedia/Wikibooks, Paracrawl, CHAVE (Santos &amp; Rocha, 2004), CBras, BrWaC (Wagner Filho et al., 2018)</td>
<td>4B</td>
</tr>
<tr>
<td>Spanish</td>
<td>EuroParl (Kohen, 2005), Literature (short stories/romances) (Garcia et al., 2019a), Wikipedia/Wikibooks</td>
<td>1.2B</td>
</tr>
<tr>
<td>English</td>
<td>EuroParl, Wikipedia/Wikibooks</td>
<td>1.6B</td>
</tr>
<tr>
<td>French</td>
<td>FrWaC (Baroni et al., 2009), Wikipedia/Wikibooks</td>
<td>2.5 B</td>
</tr>
<tr>
<td>Chinese</td>
<td>Wikipedia, Wikibooks, and literary texts</td>
<td>600M</td>
</tr>
</tbody>
</table>
Focus on collocation types with three morphosyntactic classes of bases:

- nouns
- verbs
- adjectives
### Automatically Extracted Bases

<table>
<thead>
<tr>
<th>Language</th>
<th>Nouns</th>
<th>Portuguese</th>
<th>English</th>
<th>Spanish</th>
<th>Chinese</th>
</tr>
</thead>
<tbody>
<tr>
<td>French</td>
<td>9,754</td>
<td>10,307</td>
<td>10,545</td>
<td>8,853</td>
<td>1,858</td>
</tr>
<tr>
<td>Portuguese</td>
<td>4,895</td>
<td>5,573</td>
<td>5,502</td>
<td>3,549</td>
<td>2,206</td>
</tr>
</tbody>
</table>

### Validated Candidates

<table>
<thead>
<tr>
<th>Language</th>
<th>Nouns</th>
<th>Portuguese</th>
<th>English</th>
<th>Spanish</th>
<th>Chinese</th>
</tr>
</thead>
<tbody>
<tr>
<td>French</td>
<td>8,361</td>
<td>8,690</td>
<td>8,713</td>
<td>6,576</td>
<td>103</td>
</tr>
<tr>
<td>Portuguese</td>
<td>2,902</td>
<td>3,817</td>
<td>3,982</td>
<td>2,459</td>
<td>100</td>
</tr>
</tbody>
</table>
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

<table>
<thead>
<tr>
<th>Language</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chinese</td>
<td>5,786</td>
</tr>
<tr>
<td>Spanish</td>
<td>71,131</td>
</tr>
<tr>
<td>French</td>
<td>60,447</td>
</tr>
<tr>
<td>English</td>
<td>52,826</td>
</tr>
<tr>
<td>Portuguese</td>
<td>62,968</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>253,158</strong></td>
</tr>
</tbody>
</table>
We use GDEX-inspired heuristics (Kilgarriff 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.
Automatic Translation of Collocations (Garcia et al. 2019c)
Corpus-based Methodology for an Online Multilingual Collocations Dictionary: First Steps
Adriana Orense-Ottolino, Marcos García, Maria Eugênia Olimpio de Oliveira Silva, Marie-Claude L'Hommé, Margarita Alonso Ramos, Carlos Roberto Valencia, Willam Tenório

Abstract
This paper describes the first steps of a corpus-based methodology for the development of an online multilingual collocations dictionary (PLATCOOL). 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 (Orense-Ottolino, 2017). The methodology also relies on the combination of automatic methods to extract candidate collocations (García et al., 2008) with careful post-editing performed by lexicographers. The automatic procedures take advantage of NLP tools to annotate large corpora with lemmas, part-of-speech and dependency relations in five languages (English, French, Portuguese, Spanish and Chinese). Using these data, we apply statistical measures (Bovet et al., 2017; García et al., 2009) and distributional semantic strategies to select the candidates (García et al., 2008) and retrieve corpus-based examples (Kügler et al., 2008). We also rely on automatic definition extraction (Bond & Foster, 2013) so that collocations can be more efficiently 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, 2010; Lauer, 2011; Orense-Ottolino, 2012; Turner & Berndt, 2017, among others). Despite this fact, when it comes to the teaching 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; Gregori-Gómez & Molina, 2011; Orense-Ottolino, 2009, 2012, forthcoming).
Signature

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

- collect signature
  Additional information: shouldn't there be a constant signature audit? Did the 500-3000 sample signatures have to be rechecked?

- forge signature
  Leaving personal money in a public place or on a public display to allow forgeries of signatures is used to identify an act of theft. Account for a community group within a week according to one of the city's leading handwriting. There was no one.

- gather signature
  Many city and officials should take a closer look at the signatures online. It's currently not allowed in any country, said Demata Zach, who works on elections at the National Conference of State Legislators.

- validate signature
  A week ago, I read a validation of signatures on D.C. letters that were specifically ordered as confidential in the agreement and had both signals.

- digital signature
  In California, a website sending an e-mail to legislative candidates in a state of the state's legal electronic program have pushed for digital signature collection options.

- electronic signature
  Today, there are seven states that accept electronic signatures that contain online notarizations for most public transactions and the recording of a Will of a beneficiary.

- valid signature
  Electronic signatures are valid. With a valid signature, a check can be used in conjunction with the paper trail, preventing hacking and allowing for a verifiable check that can then be easily validated.
✓ 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
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.
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.
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).

<table>
<thead>
<tr>
<th>Collocation</th>
<th>Translation</th>
<th>Trust value</th>
</tr>
</thead>
<tbody>
<tr>
<td>fazer diferença</td>
<td>faire différence</td>
<td>0.98</td>
</tr>
<tr>
<td>fazer diferença</td>
<td>hacer diferencia</td>
<td>0.90</td>
</tr>
<tr>
<td>fazer diferença</td>
<td>make difference</td>
<td>0.93</td>
</tr>
<tr>
<td>ter condição</td>
<td>devoir condition</td>
<td>0.87</td>
</tr>
<tr>
<td>ter condição</td>
<td>hacer condición</td>
<td>0.89</td>
</tr>
<tr>
<td>ter condição</td>
<td>have condition</td>
<td>0.94</td>
</tr>
<tr>
<td>ter opção</td>
<td>pouvoir option</td>
<td>0.97</td>
</tr>
<tr>
<td>ter opção</td>
<td>tener opción</td>
<td>0.96</td>
</tr>
<tr>
<td>ter opção</td>
<td>give option</td>
<td>0.72</td>
</tr>
</tbody>
</table>
Collocations Dictionary Writing System (COLDWS)

http://www.institucional.grupogbd.com/dicionario2/index
<table>
<thead>
<tr>
<th>#</th>
<th>Cod.</th>
<th>Headword</th>
<th>Part of speech</th>
<th>Status</th>
<th>Number of meanings</th>
<th>Number of collocations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>9500</td>
<td>abaissement</td>
<td>nom</td>
<td></td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>9501</td>
<td>abandon</td>
<td>nom</td>
<td></td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>9502</td>
<td>abattage</td>
<td>nom</td>
<td></td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>9503</td>
<td>abattement</td>
<td>nom</td>
<td></td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>9504</td>
<td>abattoir</td>
<td>nom</td>
<td></td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>9505</td>
<td>abbatiale</td>
<td>nom</td>
<td></td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>
### Register Headword

<table>
<thead>
<tr>
<th>Headword Information</th>
<th>Senses</th>
<th>Corresponding headwords and Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Headword language</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Headword</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Frequency</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Part of Speech</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Attention:** Please first type the headword language.

---

### Senses

**Does it have Only one sense?**

- **Yes**
- **No**

**Add sense**

---

### Corresponding Headwords

**Add Corresponding Headword**

**Comments**

**Comments about the Headword**

---

**Page 2 out of 3**

---

**Go back to previous page**

---

**Save**
Register collocation

Collocation information | Examples | Corresponding collocations, Comments and Observations

Headword meaning related to the collocation

Only one meaning

Collocation

Mutual Information
4.41289857012632

Delta pi
0.0196665241556221

SL
5424.3675641534

Log Dice
157.864151728346

LL
34.2039372944186

ZS
2804.43259119987

Synonymous collocations

Add synonymous collocation

Collocations notes

Usage notes

Lexicographical remarks
<table>
<thead>
<tr>
<th>Example 1</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Bold</strong>  <strong>Italic</strong>  <strong>Underline</strong></td>
</tr>
<tr>
<td>Por certo que esse reconhecimento não está restrito a as <strong>pessoas jurídicas</strong> de direito privado.</td>
</tr>
</tbody>
</table>

**Example source**
https://www.google.com/

<table>
<thead>
<tr>
<th>Example 2</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Bold</strong>  <strong>Italic</strong>  <strong>Underline</strong></td>
</tr>
<tr>
<td>Por outro lado, por mais redundante que possa soar, <strong>os entes públicos</strong>, enquanto <strong>pessoas jurídicas</strong>, são <strong>pessoas</strong>.</td>
</tr>
</tbody>
</table>
### 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.
### Collocations

**List of collocations**

<table>
<thead>
<tr>
<th>#</th>
<th>Cod.</th>
<th>Collocation</th>
<th>Part of speech</th>
<th>Status</th>
<th>Number of corresponding collocations</th>
<th>Related to the headword</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>20</td>
<td>amar loucamente</td>
<td>verbo</td>
<td></td>
<td>0</td>
<td>amar</td>
</tr>
<tr>
<td>2</td>
<td>2000</td>
<td>pessoas jurídicas</td>
<td>substantivo</td>
<td></td>
<td>1</td>
<td>pessoa</td>
</tr>
<tr>
<td>3</td>
<td>2002</td>
<td>fazer parte</td>
<td>substantivo</td>
<td></td>
<td>2</td>
<td>parte</td>
</tr>
<tr>
<td>4</td>
<td>2003</td>
<td>último ano</td>
<td>substantivo</td>
<td></td>
<td>3</td>
<td>ano</td>
</tr>
<tr>
<td>5</td>
<td>2004</td>
<td>grande parte</td>
<td>substantivo</td>
<td></td>
<td>2</td>
<td>parte</td>
</tr>
<tr>
<td>6</td>
<td>2006</td>
<td>ter direito</td>
<td>substantivo</td>
<td></td>
<td>4</td>
<td>direito</td>
</tr>
<tr>
<td>7</td>
<td>2007</td>
<td>entrar contato</td>
<td>substantivo</td>
<td></td>
<td>3</td>
<td>contato</td>
</tr>
<tr>
<td>8</td>
<td>2013</td>
<td>primeiro ano</td>
<td>substantivo</td>
<td></td>
<td>1</td>
<td>ano</td>
</tr>
<tr>
<td>9</td>
<td>2014</td>
<td>rede social</td>
<td>substantivo</td>
<td></td>
<td>1</td>
<td>rede</td>
</tr>
<tr>
<td>10</td>
<td>2015</td>
<td>chamar atenção</td>
<td>substantivo</td>
<td></td>
<td>4</td>
<td>atenção</td>
</tr>
</tbody>
</table>
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.

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.
song

Headword language: English
Frequency: 675400

Headword part of speech: noun

unreleased song

Language: English
Collocate: unreleased
Taxonomy: No taxonomy
Part of speech: noun
Morphosyntactic structure: No morphosyntactic structure
Delta P: 0.020625495852637
LL: 38.0714709616693

Complement:
Frequency: 1487
T-Score: 0.010790354904168
Mutual information: 6.29774623565067
Log Dice: 337.679672476587
SL: 10046.1016135151
ZS: 5545.32343757153

Related to sense
Only one meaning

Usage notes
No usage notes

Under review 🟢
**song**

**Headword language:** English  
**Frequency:** 675400

**Headword part of speech:** noun

---

**famous song**

**Language:** English  
**Collocate:** famous  
**Taxonomy:** Adjectival  
**Part of speech:** noun  
**Morphosyntactic structure:** Adjective + Noun  
**Delta Pi:** 0.0150001368625626  
**LL:** 35.6013998710351

**Complement:** No complement  
**Frequency:** 1644  
**T-Score:** 0.0109972930282427  
**Mutual Information:** 3.0355384424695  
**Log Dice:** 101.943617184775  
**Sl:** 4631.1976783903  
**ZS:** 2083.1210296154

**Related to sense**  
Only one meaning

**Usage notes**  
No usage notes

---

[Approved]
<table>
<thead>
<tr>
<th>#</th>
<th>Cod.</th>
<th>Name</th>
<th>Translation</th>
<th>Translation language</th>
<th>Type</th>
<th>Translation status</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>9003</td>
<td>sign</td>
<td>assinar</td>
<td>Portuguese</td>
<td>Headword</td>
<td></td>
</tr>
<tr>
<td></td>
<td>67001</td>
<td>organic chemistry</td>
<td>有机 化学</td>
<td>Mandarin</td>
<td>Collocation</td>
<td></td>
</tr>
<tr>
<td></td>
<td>67001</td>
<td>organic chemistry</td>
<td>orgánico química</td>
<td>Spanish</td>
<td>Collocation</td>
<td></td>
</tr>
<tr>
<td></td>
<td>67001</td>
<td>organic chemistry</td>
<td>chimie organique</td>
<td>French</td>
<td>Collocation</td>
<td></td>
</tr>
<tr>
<td></td>
<td>67003</td>
<td>organic molecule</td>
<td>orgánico molécula</td>
<td>Spanish</td>
<td>Collocation</td>
<td></td>
</tr>
<tr>
<td></td>
<td>67004</td>
<td>page talk</td>
<td>página discusión</td>
<td>Spanish</td>
<td>Collocation</td>
<td></td>
</tr>
<tr>
<td></td>
<td>67004</td>
<td>page talk</td>
<td>comentario blog</td>
<td>Portuguese</td>
<td>Collocation</td>
<td></td>
</tr>
<tr>
<td></td>
<td>67007</td>
<td>time interval</td>
<td>intervalo tiempo</td>
<td>Spanish</td>
<td>Collocation</td>
<td></td>
</tr>
<tr>
<td></td>
<td>67007</td>
<td>time interval</td>
<td>intervalo tiempo</td>
<td>Portuguese</td>
<td>Collocation</td>
<td></td>
</tr>
<tr>
<td></td>
<td>67007</td>
<td>time interval</td>
<td>temps intervalle</td>
<td>French</td>
<td>Collocation</td>
<td></td>
</tr>
</tbody>
</table>
### List of Users

<table>
<thead>
<tr>
<th>Cod.</th>
<th>Name</th>
<th>Email</th>
<th>Status of User</th>
<th>Type of User</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Adriane Stilano</td>
<td><a href="mailto:adriane.stilano@unesp.br">adriane.stilano@unesp.br</a></td>
<td>Active user</td>
<td>Main Administrator</td>
</tr>
<tr>
<td>2</td>
<td>Gislene Lira</td>
<td><a href="mailto:gislene.lira@gmail.com">gislene.lira@gmail.com</a></td>
<td>Active user</td>
<td>Main Administrator</td>
</tr>
<tr>
<td>3</td>
<td>João Pedro</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Marie-Claude L'Homme</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Maria Euginia Olimpio de Oliveira</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Margarida Maria Amares</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Marcos Garcia Gonzalez</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Antonio Pasini</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Lais Real</td>
<td><a href="mailto:lais.real@unesp.br">lais.real@unesp.br</a></td>
<td>Active user</td>
<td>Student</td>
</tr>
<tr>
<td>10</td>
<td>Emanoel Alves</td>
<td><a href="mailto:emanoel.alves@unesp.br">emanoel.alves@unesp.br</a></td>
<td>Active user</td>
<td>Student</td>
</tr>
<tr>
<td>11</td>
<td>Ariane Caldas</td>
<td><a href="mailto:ariane.caldas@unesp.br">ariane.caldas@unesp.br</a></td>
<td>Active user</td>
<td>Student</td>
</tr>
<tr>
<td>12</td>
<td>Leila Filippini</td>
<td><a href="mailto:leila.filippini@unisagrado.edu.br">leila.filippini@unisagrado.edu.br</a></td>
<td>Active user</td>
<td>Student</td>
</tr>
<tr>
<td>13</td>
<td>Daniel Meneses de Souza</td>
<td><a href="mailto:dim.souza@unesp.br">dim.souza@unesp.br</a></td>
<td>Active user</td>
<td>Main Administrator</td>
</tr>
<tr>
<td>14</td>
<td>Daniel Meneses</td>
<td><a href="mailto:danielmanesms1995@gmail.com">danielmanesms1995@gmail.com</a></td>
<td>Active user</td>
<td>Student</td>
</tr>
<tr>
<td>15</td>
<td>Luiz Carlos Alves Junior</td>
<td><a href="mailto:luiz.alvesjr.2019@gmail.com">luiz.alvesjr.2019@gmail.com</a></td>
<td>Active user</td>
<td>Student</td>
</tr>
<tr>
<td>16</td>
<td>Beatriz Martinez Rossi</td>
<td><a href="mailto:beatriz.rossi@unesp.br">beatriz.rossi@unesp.br</a></td>
<td>Active user</td>
<td>Student</td>
</tr>
</tbody>
</table>

Mostrando 11 até 18 de 18 registros.
There is also a functionality with which reviewers can post-edit entries, collocations and translations.
<table>
<thead>
<tr>
<th>ID</th>
<th>Term</th>
<th>Collocation</th>
<th>Review Status</th>
<th>Review Button</th>
</tr>
</thead>
<tbody>
<tr>
<td>99721</td>
<td>film song</td>
<td>Collocation</td>
<td>awaiting review</td>
<td>Review</td>
</tr>
<tr>
<td>100739</td>
<td>recording song</td>
<td>Collocation</td>
<td>awaiting review</td>
<td>Review</td>
</tr>
<tr>
<td>100775</td>
<td>play song</td>
<td>Collocation</td>
<td>awaiting review</td>
<td>Review</td>
</tr>
<tr>
<td>101652</td>
<td>only song</td>
<td>Collocation</td>
<td>awaiting review</td>
<td>Review</td>
</tr>
<tr>
<td>101542</td>
<td>receive song</td>
<td>Collocation</td>
<td>awaiting review</td>
<td>Review</td>
</tr>
<tr>
<td>101706</td>
<td>dance song</td>
<td>Collocation</td>
<td>awaiting review</td>
<td>Review</td>
</tr>
<tr>
<td>101946</td>
<td>feature song</td>
<td>Collocation</td>
<td>awaiting review</td>
<td>Review</td>
</tr>
<tr>
<td>102014</td>
<td>successful song</td>
<td>Collocation</td>
<td>awaiting review</td>
<td>Review</td>
</tr>
<tr>
<td>102126</td>
<td>describe song</td>
<td>Collocation</td>
<td>awaiting review</td>
<td>Review</td>
</tr>
</tbody>
</table>
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: 0.0194172172671898
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
feature song

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

SL: 2450.18767693624
ZS 1309.20836728811
famous song

Language: English

Collocation approved
The collocation was successfully approved.

Complement: No complement
SL: 4031.19767835903
ZS 2083.1210296154

Related to meaning
Only one meaning

Usage notes
No usage notes
feature song

Language: English
Collocate: feature
Taxonomy: No taxonomy
Part of speech: noun
Morphosyntactic structure: No morphosyntactic structure
Delta Pi: 0.0227624802869384
LL: 31.691804015703

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

We aim to objectively establish an appropriate methodology and phraseographic model for PLATCOL.
Finish developing the end-user platform in order to launch PLATCOL until the end of 2022!

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.
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).
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?