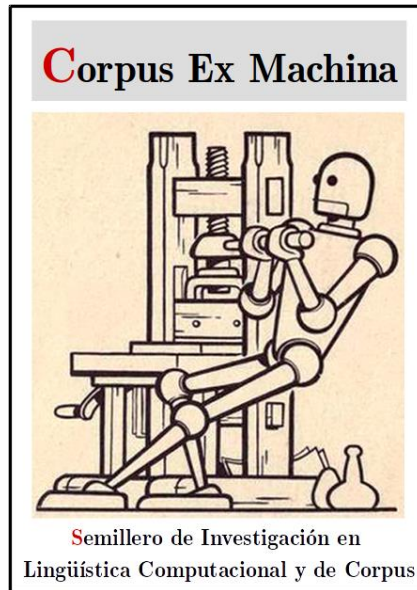


# *Corpus Ex Machina*



## **Semillero de Investigación en Lingüística Computacional y de Corpus**

### **Descripción**

*Corpus Ex Machina* es un semillero de investigación en las áreas de la Lingüística Computacional y de la Lingüística de Corpus, que forma parte del GES (Grupo de Estudios Sociolingüísticos), adscrito a la Facultad de Comunicaciones de la Universidad de Antioquia. El semillero inicia labores a comienzos del año 2016 como resultado del trabajo realizado en la asignatura de Lingüística Computacional del pregrado Letras: Filología Hispánica. Hasta este momento, *Corpus Ex Machina* ha contado con la participación de estudiantes de varias Facultades de la Universidad de Antioquia (estudiantes y egresados de los pregrados Letras: Filología Hispánica, Filología Hispánica y de la Licenciatura en Educación Básica con énfasis en Humanidades, Lengua Castellana). Además, desde el año 2017 el semillero se encuentra adscrito a RedSin (Red de Semilleros de la Universidad de Antioquia).

Dentro de su espíritu pluridisciplinario, el semillero se encuentra abierto a la comunidad universitaria en general; proponiendo la posibilidad de trabajar con estudiantes adscritos a diferentes dependencias que posean áreas afines con nuestros proyectos.

### **Objetivos**

*Corpus Ex Machina* tiene como objetivo principal afianzar el conocimiento en el área de la Lingüística Computacional y trabajar sobre corpus lingüísticos. Busca, para ello, que los estudiantes se formen en el campo de la investigación, utilizando los procedimientos y las herramientas que la informática provee para el análisis de dichos corpus.

Este semillero tiene además como objetivo, el fundamentarse en la interdisciplinariedad como forma de trabajo, ya que consideramos que a partir de esta podremos avanzar como investigadores mejor formados y complementados.

## Áreas de trabajo y líneas de investigación

- ❖ LINGÜÍSTICA COMPUTACIONAL.
- ❖ LINGÜÍSTICA DE CORPUS.
- ❖ MODELIZACIÓN DE SISTEMAS QUE FUNCIONAN A PARTIR DE HERRAMIENTAS QUE TRATAN EL LENGUAJE NATURAL.
- ❖ PROCESAMIENTO DEL LENGUAJE NATURAL.
- ❖ SEMÁNTICA DE CORPUS.
- ❖ SISTEMAS PARA EL APRENDIZAJE DE LENGUAS ASISTIDOS POR ORDENADOR (ALAO).
- ❖ ESPAÑOL LENGUA EXTRANJERA (ELE).
- ❖ RELACIÓN ENTRE LINGÜÍSTICA Y MÚSICA.

## Coordinadores del semillero

JORGE MAURICIO MOLINA MEJÍA

Doctor en Informática y Ciencias del Lenguaje.

Docente de cátedra Facultad de Comunicaciones Universidad de Antioquia.

Programa Letras: Filología Hispánica.

Contacto: [jorge.molina@udea.edu.co](mailto:jorge.molina@udea.edu.co)

LAURA MARCELA QUINTERO MONTOYA

Magíster en Lingüística Aplicada.

Docente de cátedra Facultad de Comunicaciones Universidad de Antioquia

Programa Cátedra Universitaria en Lengua Materna.

Contacto: [lmargela.quintero@udea.edu.co](mailto:lmargela.quintero@udea.edu.co)

## Proyectos de investigación en los que participa *Corpus Ex Machina*

- ❖ Proyecto **Preseca-Medellín**, etiquetado morfosintáctico del corpus oral de hablantes del Valle de Aburrá. Proyecto terminado en febrero de 2019.
- ❖ Proyecto **DICEELE** (Dispositivo Informático basado en Corpus para la Enseñanza del

Español Lengua Extranjera). Proyecto actualmente en ejecución.

- ❖ Proyecto **CODITEX** (Corpus de Artículos Científicos de Comunidades Discursivas). Proyecto terminado en 2017, trabajo de grado para optar al título de Filóloga Hispanista de Mariana Múnera Manco.
- ❖ Proyecto **MusaLingua** (Corpus para el análisis Lingüístico de la Música). Proyecto actualmente en ejecución.

## Trabajos de grado realizados en el semillero

- ❖ Asesoría de trabajo de grado en el pregrado en Letras: Filología Hispánica de la estudiante **Mariana Múnera Manco**, Facultad de Comunicaciones, Universidad de Antioquia, Medellín (Colombia). Sustentación pública ante jurado el 6 de diciembre de 2017. Miembros del jurado: Dr. Jorge Mauricio Molina Mejía, docente-investigador Universidad de Antioquia (Asesor del monográfico) y Dr. Gerardo Eugenio Sierra Martínez, profesor Universidad Nacional Autónoma de México (Evaluador). Título del trabajo de grado: *"Identificación y análisis de contextos definitorios sintácticos: clasificación tipológica definicional a partir de un corpus de artículos científicos"*.
- ❖ Asesoría de trabajo de grado en el pregrado en Letras: Filología Hispánica del estudiante **David Alexis Arboleda Méndez**, Facultad de Comunicaciones, Universidad de Antioquia, Medellín (Colombia). Sustentación pública ante jurado el 12 de junio de 2018. Miembros del jurado: Dr. Jorge Mauricio Molina Mejía, docente-investigador Universidad de Antioquia (Asesor del monográfico) y Mg. María Isabel Marín Morales, profesora Tecnológico de Antioquia, Institución universitaria (Evaluadora). Título del trabajo de grado: *"Propuesta para un modelo de etiquetado y análisis de la negación sintáctica en español basado en gramática de dependencias"*.
- ❖ Asesoría de trabajo de grado en el pregrado Filología Hispánica del estudiante **Andrés Felipe Grajales Ramírez**, Facultad de Comunicaciones, Universidad de Antioquia, Medellín (Colombia). Sustentación pública ante jurado el 23 de mayo de 2019. Miembros del jurado: Dr. Jorge Mauricio Molina Mejía, docente-investigador Universidad de Antioquia (Asesor del monográfico) y Mg. Laura Marcela Quintero Montoya, profesora Universidad de Antioquia (Evaluadora). Título del trabajo de grado: *"Problemas del tratamiento computacional de la anáfora: Análisis en un corpus para la enseñanza del Español Lengua Extranjera"*.
- ❖ Asesoría de trabajo de grado en el pregrado Letras: Filología Hispánica del estudiante **Arbey Stiven Osorio Osorio**, Facultad de Comunicaciones, Universidad de Antioquia, Medellín (Colombia). Sustentación pública ante jurado el 23 de mayo de 2019. Miembros del jurado: Dr. Jorge Mauricio Molina Mejía, docente-investigador Universidad de Antioquia (Asesor del monográfico) y Mg. Laura Marcela Quintero Montoya, profesora Universidad de Antioquia (Evaluadora). Tema del trabajo de grado: *"Estudio exploratorio del error gramatical en la comunicación escrita de aprendientes de ELE en la ciudad de Medellín"*.

- ❖ Asesoría de trabajo de grado en el pregrado en Filología Hispánica de la estudiante **Sara María Galvis Amaya**, Facultad de Comunicaciones, Universidad de Antioquia, Medellín (Colombia). Título del trabajo de grado: "*Hacia un análisis de la correferencia textual en español y en francés a partir de un corpus de textos literarios*". En curso.
- ❖ Asesoría de trabajo de grado en el pregrado en Letras: Filología Hispánica del estudiante **José Luis Pemberty Tamayo**, Facultad de Comunicaciones, Universidad de Antioquia, Medellín (Colombia). Título del trabajo de grado: "*Hacia la elaboración de un sistema de etiquetado morfosintáctico semiautomático para lenguas minoritarias*". En curso.

## **Producción científica y académica de *Corpus Ex Machina***

### **Artículos publicados:**

- MOLINA MEJÍA, Jorge Mauricio (2019). «Hacia un análisis de la obra de Richard Wagner a través de la lingüística computacional». In *Forma y Función*. Vol. 32, num 1, pp 125-147. DOI: 10.15446/fyf.v32n1.77419. ISSN electrónico: 2256-5469. URL: <https://revistas.unal.edu.co/index.php/formayfuncion/article/view/77419>.
- MOLINA MEJÍA, Jorge Mauricio; GRAJALES RAMÍREZ, Andrés Felipe & PEMBERTY TAMAYO, José Luis (2019). «Hacia un dispositivo informático basado en Corpus para la Enseñanza del Español Lengua Extranjera (DICEELE)». In *Revista Internacional de Tecnología, Conocimiento y Sociedad*. Vol. 7, num 1, 2019, pp 1-13. DOI: 10.18848/2474-588X/CGP/v07i01/1-13. ISSN: 2174-8985.
- MOLINA MEJÍA, Jorge Mauricio; GRAJALES RAMÍREZ, Andrés Felipe; PEMBERTY TAMAYO, José Luis & ZAPATA GRANADOS, María Adelaida (2018). «Corpus Ex Machina: Computación y lengua española». In *Experimenta: Revista de divulgación científica de la Universidad de Antioquia*. Edición 9, enero-junio 2018, pp 52-55. ISSN: 2357-3503.
- MOLINA MEJÍA, Jorge Mauricio (2017). «Des corpus annotés au coeur d'une plate-forme, pour la formation linguistique des enseignants de FLE en Colombie». In *CORELA: Cognition, représentation, langage*. [En ligne], HS-21 | 2017 : Linguistique de corpus : vues sur la constitution, l'analyse et l'outillage (Sous la direction de Iris ESHKOL-TARAVELLA et Anaïs LEFEUVRE-HALFTERMEYER), puesto en línea el 02 de febrero de 2017, publicado por el CerLiCO (Cercle Linguistique du Centre et de l'Ouest). URL: <http://corela.revues.org/4857>; ISSN electrónico: 1638-573X.
- MOLINA MEJÍA, Jorge Mauricio (2016). «Lingüística de corpus utilizada en la concepción de plataformas de formación de docentes de lenguas». In *CHIMERA: Romance Corpora and Linguistic Studies*. [En línea], Vol. 3, n° 1 | 2016, pp. 57-85, puesto en línea el 14 de diciembre 2016, Portal de Revistas Electrónicas, Universidad Autónoma de Madrid. URL: <https://revistas.uam.es/index.php/chimera/article/view/5564> DOI: 10.15366/chimera. ISSN electrónico: 2386-2629.

## **Ponencias presentadas:**

- GRAJALES RAMÍREZ, Andrés Felipe & ZAPATA GRANADOS, María Adelaida (2019). El tratamiento computacional de la anáfora en un dispositivo informático para la enseñanza del Español Lengua Extranjera. Presentado en el Encuentro de Semilleros de Investigación – ESI 2019. Universidad de Antioquia, Medellín – Colombia. 18 y 19 de marzo de 2019. Presentado el 19 de marzo de 2018.
- MUNERA MANCO, Mariana & MOLINA MEJÍA, Jorge Mauricio (2018). *Hacia una mirada específica en el conocimiento terminográfico: análisis de contextos definitorios sintácticos en producciones científicas*. Presentado en XVI Simposio de la Red Iberoamericana de Terminología. Universidad Autónoma de Manizales, Manizales - Colombia. Presentado el 23 de octubre de 2018.
- MOLINA MEJÍA, Jorge Mauricio & GONZÁLEZ RÁTIVA, María Claudia (2018). *Etiquetado morfosintáctico y modelizado a partir del corpus lingüístico Preseaa-Medellín*. Presentado en XI Seminario Internacional en Ciencias de la Computación (SICC 2018) y II Congreso Internacional de Lingüística Computacional y de Corpus (CILCC 2018). Universidad de Medellín, Medellín - Colombia. Presentado el 12 de octubre de 2018.
- MARÍN MORALES, María Isabel; MOLINA MEJÍA, Jorge Mauricio & PEMBERTY TAMAYO, José Luis (2018). *Etiquetado lingüístico de un corpus para la enseñanza del español como lengua extranjera: enfoque metodológico*. Presentado en XI Seminario Internacional en Ciencias de la Computación (SICC 2018) y II Congreso Internacional de Lingüística Computacional y de Corpus (CILCC 2018). Universidad de Medellín, Medellín - Colombia. Presentado el 11 de octubre de 2018.
- MOLINA MEJÍA, Jorge Mauricio & ANTONIADIS, Georges (2018). *Sistemas de ALAO basados en corpus y diseñados a partir de herramientas de Lingüística Computacional*. Presentado en XI Seminario Internacional en Ciencias de la Computación (SICC 2018) y II Congreso Internacional de Lingüística Computacional y de Corpus (CILCC 2018). Universidad de Medellín, Medellín - Colombia. Presentado el 11 de octubre de 2018.
- MOLINA MEJÍA, Jorge Mauricio; ZAPATA GRANADOS, María Adelaida; TABORDA OBANDO, Daniel & OSORIO OSORIO, Arbey Stiven (2018). *Concepción de un corpus especializado para la enseñanza de ELE*. Presentado en X Coloquio Internacional sobre Investigación en Lenguas Extranjeras. Universidad Veracruzana, Veracruz - Mexico. Presentado el 21 de junio de 2018.
- MOLINA MEJÍA, Jorge Mauricio & ROCHA TORRES, Karen Paola (2018). *Hacia la elaboración de un corpus especializado para la enseñanza del Español Lengua Extranjera*. Presentado en Encuentro Internacional del Grupo de Estudios Sociolingüísticos - GES 2018, Grupo de Estudios Sociolingüísticos. Facultad de Comunicaciones – Universidad de Antioquia, Medellín - Colombia. Presentado el 11 de mayo de 2018.
- MÚNERA MANCO, Mariana & MOLINA MEJÍA, Jorge Mauricio (2018). *Las comunidades discursivas en Colombia: Estudio terminográfico de un corpus de artículos científicos*. Presentado en V Conferencia Internacional sobre Contenido Multilingüe y Comunicación

Especializada para la Industria y el Comercio; y X Seminario Nacional de Terminología. Universidad de Antioquia, Medellín - Colombia. Presentado el 24 de marzo de 2018.

ZAPATA GRANADOS, María Adelaida & PEMBERTY TAMAYO, José Luis. (2018). *Dispositivo Informático basado en Corpus para la Enseñanza del Español Lengua Extranjera (DICEELE)*. Evento: Encuentro de Semilleros de Investigación – ESI 2018. Universidad de Antioquia, Medellín – Colombia. 22 y 23 de marzo de 2018. Presentado el 23 de marzo de 2018.

MOLINA MEJÍA, Jorge Mauricio (2017). *Lingüística de corpus y enseñanza de lenguas extranjeras: herramientas y etiquetado de textos*. Presentado en The First Fall School of Corpus and Computational Linguistics, Grupo en Traducción y Nuevas Tecnologías, Escuela de Idiomas – Universidad de Antioquia. Medellín - Colombia. Presentado el 13 de diciembre de 2017.

MOLINA MEJÍA, Jorge Mauricio (2017). *Linguistique informatique et son application dans la didactique des langues et dans les études littéraires et humanistes*. Presentado en Première Rencontre des Chercheurs Franco-Colombiens, Universidad Externado de Colombia y Ambassade de France en Colombie. Bogotá - Colombia. Presentado el 05 de diciembre de 2017.

MÚNERA MANCO, Mariana (2017). *Hacia un desarrollo en el conocimiento terminográfico de las comunidades discursivas: identificación de contextos definitorios sintácticos*. Presentado en Coloquio de Lingüística, del Grupo de Estudios Sociolingüísticos – GES, Facultad de Comunicaciones, Universidad de Antioquia, Medellín – Colombia. Presentado el 29 de noviembre de 2017.

MOLINA MEJÍA, Jorge Mauricio (2017). *La Lingüística Computacional y su aplicación en la didáctica de lenguas y en los estudios literarios y humanísticos*. Presentado en III Jornadas de Estudiantes de Lingüística y Literatura, Facultad de Comunicaciones - Universidad de Antioquia. Medellín - Colombia. Presentado el 27 de octubre de 2017.

MÚNERA MANCO, Mariana (2017). *Identificación y análisis de contextos definitorios sintácticos: clasificación tipológica definicional a partir de un corpus de artículos científicos*. Presentado en III Jornadas de Estudiantes de Lingüística y Literatura, Facultad de Comunicaciones - Universidad de Antioquia. Medellín - Colombia. Presentado el 26 de octubre de 2017.

GRAJALES RAMÍREZ, Andrés Felipe & PEMBERTY TAMAYO, José Luis (2017). *DICEELE: Hacia la concepción de un dispositivo informático basado en corpus para la enseñanza del Español como Lengua Extranjera*. Presentado en Coloquio de Lingüística, del Grupo de Estudios Sociolingüísticos – GES, Facultad de Comunicaciones, Universidad de Antioquia, Medellín – Colombia. Presentado el 12 de octubre de 2017.

MOLINA MEJÍA, Jorge Mauricio; PEMBERTY TAMAYO, José Luis; GRAJALES RAMÍREZ, Andrés Felipe & BERMÚDEZ CARDONA, Alejandra (2017). *DICEELE: Dispositivo Informático Basado en Corpus para la Enseñanza de ELE*. Presentado en “IX Coloquio Internacional sobre Investigación en Lenguas Extranjeras” CIILE 2017. Facultad de Ciencias de la Educación, Universidad de Granada. Granada - España. Presentado el 23 de junio 2017.

GRAJALES RAMÍREZ, Andrés Felipe & PEMBERTY TAMAYO, José Luis (2017). *Dispositivo*

*Informático basado en Corpus para la Enseñanza del Español Lengua Extranjera (DICEELE)*. Presentado en XVI Encuentro Departamental de Semilleros de Investigación. Universidad de Antioquia, Medellín – Colombia. Presentado el 20 de mayo de 2017.

MOLINA MEJÍA, Jorge Mauricio; GONZÁLEZ RÁTIVA, María Claudia; PEMBERTY TAMAYO, José Luis; GRAJALES RAMÍREZ, Andrés Felipe & BERMÚDEZ CARDONA, Alejandra (2017). *Hacia la anotación y etiquetado de un corpus sociolingüístico: Presea-Medellín*. Presentado en “Congreso Internacional de Lingüística Computacional y de Corpus”. Instituto Caro y Cuervo. Bogotá - Colombia. Presentado el 19 de mayo 2017.

MOLINA MEJÍA, Jorge Mauricio; GRAJALES RAMÍREZ, Andrés Felipe; PEMBERTY TAMAYO, José Luis; ARANGO RIVERA, Daniel & BARRERA ORTEGA, Juan Esteban (2017). *Hacia la constitución de un dispositivo informático basado en corpus para la enseñanza del español como lengua extranjera*. Presentado en “Congreso Internacional de Lingüística Computacional y de Corpus. Instituto Caro y Cuervo”. Bogotá - Colombia. Presentado el 19 de mayo 2017.

GRAJALES RAMÍREZ, Andrés Felipe & PEMBERTY TAMAYO, José Luis (2017). *Dispositivo Informático basado en Corpus para la Enseñanza del Español Lengua Extranjera (DICEELE)*. Presentado en Encuentro de Semilleros de Investigación – ESI 2017. Universidad de Antioquia, Medellín – Colombia. Presentado el 05 de mayo de 2017.

MOLINA MEJÍA, Jorge Mauricio & ANTONIADIS, Georges (2017). *TICE et Linguistique Textuelle utilisés dans la formation en écriture des futurs enseignants de FLE : La plate-forme ELiTe-[FLE]<sup>2</sup>*. Presentado en “WRAB 2017: Writing Research Across Borders IV”. Pontificia Universidad Javeriana. Bogotá - Colombia. Presentado el 18 de febrero 2017.

PEMBERTY TAMAYO, José Luis; GRAJALES RAMÍREZ, Andrés Felipe; MOLINA MEJÍA, Jorge Mauricio; ARANGO RIVERA, Daniel & ARIAS BEDOYA, Franklin Yessid (2016). *Hacia el etiquetado de corpus orales y escritos en español: Estudio de analizadores morfosintácticos*. Presentado en “II Congreso Internacional de Investigación Lingüística y III Coloquio Internacional de Argots”. Grupo GELIR, Universidad de Antioquia. Medellín - Colombia. Presentado el 24 de noviembre 2016.

MOLINA MEJÍA, Jorge Mauricio & ANTONIADIS, Georges (2016). *Utilización de corpus etiquetados en la concepción de dispositivos informáticos para la enseñanza de lenguas*. Presentado en “II Congreso Internacional de Investigación Lingüística y III Coloquio Internacional de Argots”. Grupo GELIR, Universidad de Antioquia. Medellín - Colombia. Presentado el 24 de noviembre 2016.

MOLINA MEJÍA, Jorge Mauricio; ARANGO RIVERA, Daniel; ARIAS BEDOYA, Franklin Yessid; ARBOLEDA MÉNDEZ, David Alexis; GRAJALES RAMÍREZ, Andrés Felipe & PEMBERTY TAMAYO, José Luis (2016). *Anotación de corpus para la concepción de sistemas informáticos en enseñanza de lenguas y para la exploración de fenómenos lingüísticos: Corpus Ex Machina*. Presentado en “Primer Coloquio Colombiano de Lingüística Computacional y de Corpus”. Universidad del Valle. Cali - Colombia. Presentado el 13 de octubre 2016.