

EVOLUCIÓN LINGÜÍSTICA *IN SILICO*

De datos a gran escala a agentes artificiales que crean lenguas desde cero

Thomas Brochhagen

Todos hablamos una lengua y tenemos cierta intuición sobre ella: desde el vocabulario hasta la forma de juntar las palabras de acuerdo con su gramática. Sin embargo, todavía nos falta mucho por entender acerca de los procesos que hacen posible el lenguaje y dan forma a su evolución. Los avances computacionales recientes nos han permitido abordar estas cuestiones desde nuevos ángulos. Este artículo destaca los métodos y descubrimientos que ha traído la era de la computación, desde el aprendizaje a partir de datos a gran escala provenientes de miles de idiomas hasta la evolución de lenguas creadas por la inteligencia artificial.

Palabras clave: **lenguaje, evolución, inteligencia artificial, tipología, universales.**

A diferencia de los especímenes biológicos, las lenguas no dejan huella en el registro fósil. Esto hace que estudiar su evolución sea una tarea complicada. No tenemos pruebas directas sobre cómo se estructuraba el habla de nuestros antepasados ni sabemos qué cambios sufrió hasta su forma actual, miles de años después. La reacción más tristemente célebre ante estas dificultades fue la decisión de la Sociedad Lingüística de París en 1866 de prohibir cualquier debate sobre la cuestión (Corballis, 2008).

Pese a los obstáculos, la pregunta de cómo evolucionan los idiomas y lo que esto nos dice sobre nosotros mismos siempre ha fascinado a los expertos. La investigación no solo ha continuado hasta nuestros días, sino que ha empezado a acelerarse notablemente en la era de la computación. En este artículo, revisaremos una serie de métodos y descubrimientos facilitados por el avance de la computación.

Sin especímenes lingüísticos originales que diseccionar, la investigación sobre la evolución de las lenguas se ha centrado tradicionalmente en tres tipos de cuestiones. En primer lugar, se pueden inferir patrones regulares a

partir de registros históricos de lenguas con tradición escrita. En segundo lugar, para abordar cuestiones como si el lenguaje surgió gradual o súbitamente, se estudian las capacidades cognitivas de los humanos modernos y se comparan con las que podrían haber tenido nuestros antepasados, o con las de nuestros parientes más cercanos, como los chimpancés y los gibones. En tercer lugar, se recopilan y comparan datos sobre los diferentes idiomas hablados actualmente en todo el mundo. De esta forma, la diversidad de lenguas del presente nos sirve de ventana a los procesos evolutivos de los que son el resultado. Puesto que las lenguas modernas son la expresión de las trayectorias evolutivas de su pasado, sus puntos en común y sus diferencias pueden darnos pistas importantes.

Estas tres líneas de investigación continúan en marcha en la actualidad. No obstante, la era actual de computación accesible y asequible ha añadido nuevas capacidades y dimensiones al estudio de la evolución de las lenguas. Por un lado, los mencionados esfuerzos cuentan ahora con el apoyo de potentes algoritmos que nos permiten cuantificar las pruebas de teorías rivales sobre la dicha

**«La era actual de computación
accesible y asequible
ha añadido nuevas capacidades
y dimensiones al estudio de la
evolución de las lenguas»**

CÓMO CITAR ESTE ARTÍCULO:

Brochhagen, T. (2024). Evolución lingüística *in silico*. De datos a gran escala a agentes artificiales que crean lenguas desde cero. *Metode Science Studies Journal*. <https://doi.org/10.7203/metode.15.27692>

evolución, construir mejores mapas genealógicos de las lenguas y predecir mejor los cambios futuros. Por otro, ahora la evolución de las lenguas también se estudia artificialmente, con inteligencias artificiales que crean lenguas por sí mismas.

■ CÓMO EVOLUCIONA EL SIGNIFICADO A TRAVÉS DE LAS LENGUAS, LAS CULTURAS Y EL TIEMPO

Cada lengua tiene su idiosincrasia. En inglés, los conceptos de dedos de las manos y de los pies se expresan con dos palabras diferentes: *finger* y *toe*. En castellano, ambos remiten al mismo término: *dedo*. Sorprendentemente, estos dos conceptos también se coexpresan en más de 130 idiomas de todo el mundo (Rzyski et al., 2020). Es decir, más de 130 idiomas utilizan una sola palabra para expresar ambos conceptos: desde la lengua secoya, hablada en el Amazonas entre Ecuador y Colombia, hasta la lengua takia de Papúa Nueva Guinea. Encontramos muchos patrones similares en los vocabularios de todo el mundo. *Hoja* y *pluma* también se suelen expresar con la misma palabra, al igual que *bueno* y *bonito*, *pequeño* y *joven*, y *agujero* y *cueva*. Estos patrones no se pueden explicar por una ascendencia compartida o por proximidad geográfica. Lo que resulta necesario explicar (mediante el estudio de la evolución del lenguaje) es, por tanto, qué hay en la relación entre los dedos de las manos y de los pies, o entre *agujero* y *cueva*, o entre *hoja* y *pluma*, para que estos conceptos se asimilen en tantos idiomas. Al abordar estas cuestiones, lo que más nos interesa es qué nos dicen sobre cómo organizamos el significado los seres humanos.

Poder abordar esta cuestión basándonos no solo en un puñado de lenguas sino en datos de cientos o miles de ellas es uno de los avances recientes más importantes en este campo. Algunos ejemplos de nuevos recursos a gran escala incluyen la Base de Datos de Colexificaciones Interlingüísticas (Rzyski et al., 2020), que registra la expresión de conceptos en más de 3.000 idiomas; o Kinbank, la base de datos mundial sobre parentesco terminológico (Passmore et al., 2023), que recopila datos sobre cómo se expresan los términos de parentesco (por ejemplo, si existe un término para el concepto de *tío*, o si las ideas de *abuelo paterno* o *materno* se expresan con diferentes palabras). Conviene dejar claro que, aunque estos recursos son nuevos, se basan en el trabajo de campo fundamental de los lingüistas que han documentado –y continúan documentando– las lenguas de todo el mundo. La novedad es que este tipo de datos está ahora convenientemente digitalizado y accesible en formatos unificados, y que ahora tenemos la potencia y los métodos informáticos necesarios para procesarlos automáticamente.



Imagen de Freepik

Cada lengua tiene su idiosincrasia. En muchos idiomas se usa una misma palabra para expresar dos conceptos diferentes, como *bueno* y *bonito*, *pequeño* y *joven*, o *agujero* y *cueva*. Este patrón no se puede explicar por una ascendencia compartida o por proximidad geográfica.

«El significado se organiza de forma regular en todos los idiomas humanos, y la evolución de todos ellos sigue patrones predecibles»

Si queremos saber qué hace que algunos conceptos sean más comúnmente expresados con el mismo término que otros, el siguiente paso es captar las relaciones entre ellos. Para que esto funcione, los conceptos han de representarse de alguna manera. Esto se puede lograr aprovechando técnicas computacionales modernas. Por ejemplo, es probable que el vínculo entre *finger* y *toe* en inglés, como el de *agujero* y *cueva*, se base, al menos en cierta medida, en su similitud visual. Podemos accionar esta idea gracias a modelos modernos de visión, que procesan grandes cantidades de imágenes para construir una representación informática de ellas. De esta forma, podemos medir lo parecidos que son visualmente los dedos de las manos y de los pies. De forma parecida, podemos utilizar otros recursos para conocer, por ejemplo, cómo de similar es el contexto de uso de diferentes conceptos, su cercanía en la memoria asociativa, etc. En pocas palabras, diferentes técnicas modernas –construidas sobre el trabajo interdisciplinario en psicología, inteligencia artificial, procesamiento del lenguaje natural y la estadística– nos pueden ayudar a representar computacionalmente las diferentes formas en las que dos conceptos se asemejan (o bien qué los separa).

El resultado es un recurso que nos dice cómo se expresan los conceptos en diferentes idiomas y qué relación los une. Es decir, los dedos de manos y pies son visualmente similares, aparecen en contextos parecidos y están estre-

chamente relacionados (si oímos *finger*, podemos pensar automáticamente en la palabra *toe*). En cambio, aunque un dedo puede ser visualmente similar a una salchicha, no aparecen en el mismo contexto ni guardan una relación estrecha. Gracias a este tipo de información, recientemente se han realizado importantes descubrimientos. En primer lugar, las palabras para conceptos similares se atraen universalmente entre sí (Xu et al., 2020). Es decir, es más fácil que estos conceptos se expresen con el mismo término, ya sea en mandarín o en holandés. Es probable que esto ocurra porque expresar conceptos parecidos con la misma palabra facilita su aprendizaje. En segundo lugar, esta tendencia universal tiene un límite: dos conceptos tan parecidos que pueden confundirse no se atraen entre sí (Brochhagen y Boleda, 2022). Por ejemplo, utilizar la misma palabra para designar el *jueves* y el *miércoles* es una idea con poco futuro. Sin embargo, los términos *finger* y *toe* funcionan en inglés porque el contexto suele dejar claro a cuál nos referimos. En tercer lugar, estos patrones pueden deberse a una tendencia universal de las lenguas por resultar a la vez simples («utilizar el menor número de palabras posible») y efectivas («utilizar diferentes palabras para conceptos que queremos distinguir», como *martes* y *jueves*). En otras palabras, los idiomas están moldeados por la necesidad de mantener un equilibrio entre ser simples y ser informativos (véase Kemp y Regier, 2012; Zaslavsky et al., 2018). Si un idioma es demasiado simple, no es útil para hablar con otros. Si es demasiado complejo, es poco manejable o imposible de aprender. La evolución de las lenguas alcanza un equilibrio entre ambos extremos y explica cómo se organiza el significado en diferentes idiomas. Por último, los mismos factores que predicen si dos conceptos se atraen entre sí en diferentes idiomas explican el uso del lenguaje de los niños y las niñas (Brochhagen et al., 2023): cuando de pequeños llamamos *coche* a un barco, *perro* a una vaca o *sol* a una lámpara, estamos utilizando la misma relación semántica que se refleja universalmente en nuestras lenguas. Cuando a una criatura le falta una palabra y utiliza otra para hacerse entender (como al llamar *perrito* a una vaca), a un nivel fundamental, está haciendo algo similar a cuando, en castellano o checheno, los hablantes utilizan la misma palabra para referirse a los dedos de las manos y de los pies.

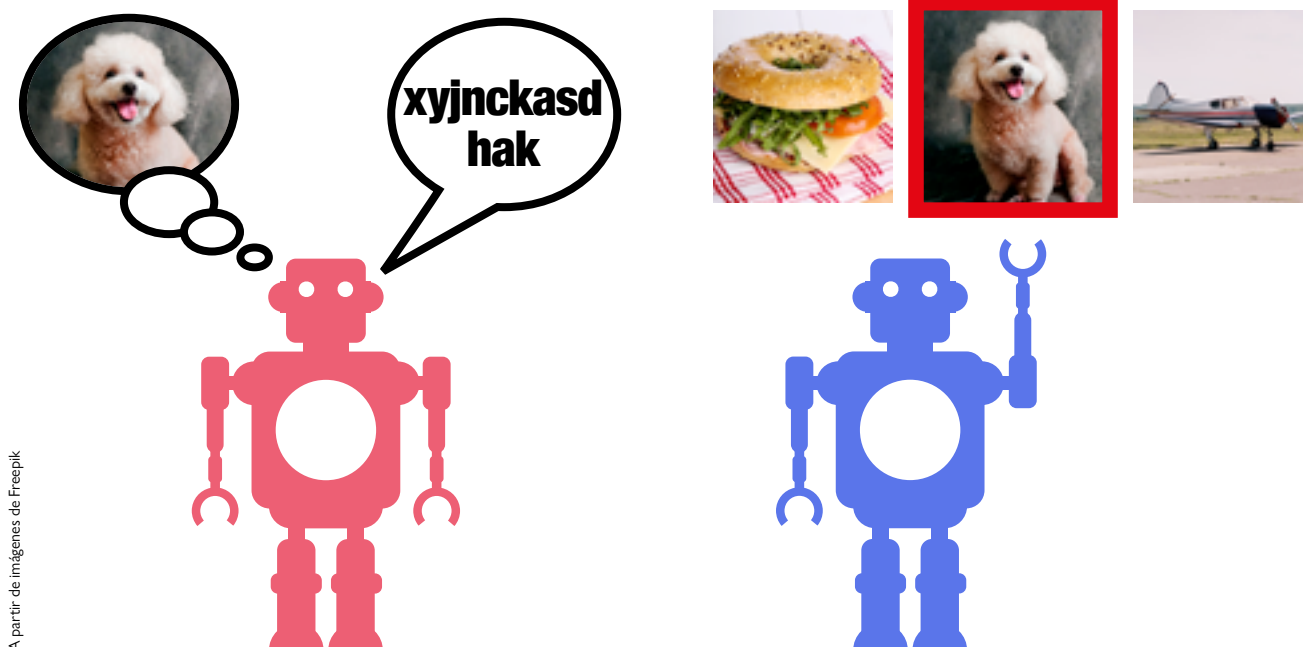
En resumen, esto indica que el significado se organiza de forma regular en todos los idiomas humanos, y que la evolución de todos ellos sigue patrones predecibles que se pueden destilar a partir de los datos. Estos patrones son el resultado de las fuerzas que dan forma al lenguaje. En particular, el impulso hacia lenguas que sean a la vez simples e informativas: lo suficientemente simples para poder aprenderlas, pero lo suficientemente informativas para que podamos entendernos los unos a los otros.

■ AGENTES ARTIFICIALES QUE INVENTAN LENGUAS DESDE CERO

Una forma radicalmente distinta de abordar la evolución de las lenguas es a través de la lente de la inteligencia artificial. Existen dos motivaciones principales para estudiar la evolución del lenguaje *in silico*. En primer lugar, las lenguas humanas son producto de nuestra biología, ecología y cultura. La forma en que procesamos, percibimos e interactuamos con el mundo es lo que determina en última instancia las propiedades del lenguaje humano. Sin embargo, existe un gran debate sobre cuáles son los factores biológicos, ecológicos o culturales responsables de propiedades lingüísticas concretas. Estudiar los agentes artificiales y su lenguaje es, por tanto, un campo de exploración prometedor. A diferencia de los humanos, en este caso controlamos hasta el último detalle de la «biología», «ecología» y «cultura» de estos entes. En segundo lugar, aunque la inteligencia artificial ha logrado avances impresionantes en los dos últimos años —sobre todo con el auge de ChatGPT y herramientas similares— las lenguas artificiales continúan muy alejadas de las humanas. Carecen de la flexibilidad y libertad de nuestro discurso. Mediante el estudio de la evolución del lenguaje artificial, esperamos aprender sobre el lenguaje, pero también cómo lograr que los agentes artificiales creen mejores lenguajes, más parecidos a los humanos (Lazaridou y Baroni, 2020).

Una configuración popular para animar a los agentes artificiales a crear sus propios lenguajes se denomina *juego de referencia*, del cual existen muchas variantes. La configuración más sencilla consiste en un emisor y un receptor. La tarea del emisor consiste en hacer que el receptor elija una imagen determinada de entre un conjunto de imágenes candidatas. Por ejemplo, tal como se ilustra en la Figura 1 en la página siguiente, la tarea puede consistir en transmitir la imagen con el perro blanco. A continuación, el emisor envía un mensaje de su creación al receptor. El receptor la interpreta y selecciona una imagen (por ejemplo, una de las tres de la Figura 1). Entonces, ambos reciben información sobre si la elección era correcta o no. Después, el juego se repite durante muchas más rondas.

¿Cómo puede surgir el lenguaje a partir de este tipo de situación? Al principio, el emisor no tiene una forma establecida de transmitir nada al receptor. Lo mejor que puede hacer es enviar un mensaje aleatorio. De igual forma, el receptor no tiene manera de interpretar el mensaje y se ve obligado a elegir la imagen al azar. No se transfiere ninguna información. Sin embargo, con el tiempo, los mensajes inicialmente aleatorios adquieren significado mediante la repetición: si el receptor adivina la imagen, es más probable que en el futuro los agentes utilicen el mismo mensaje para ese mismo tipo de imagen. Esta situación recuerda a jugar muchas rondas



A partir de imágenes de Freepik

Figura 1. Dos agentes artificiales hablando sobre imágenes de ImageNet (Deng et al., 2009). El emisor (izquierda) quiere que el receptor (derecha) seleccione una imagen (el perro blanco). Para ello, envía un mensaje. A continuación, el receptor adivina la imagen que tenía en mente el remitente, y el juego se repite con un conjunto diferente de imágenes. En la mayoría de casos, con el tiempo los agentes inventan lenguajes que les permiten comunicarse con gran precisión.

de «charadas»: la mímica que ha funcionado anteriormente se reutiliza, y esto hace el juego mucho más fluido. Si jugamos habitualmente con amigos, es probable que desarrollemos un vocabulario gestual complejo que nos permita comunicarnos rápido y bien. Este mismo principio se aplica a los agentes artificiales.

Lo interesante no es solo que los agentes sean capaces de desarrollar su propio lenguaje *ex nihilo*, sino qué tipo de lenguajes inventan. Los resultados en este campo han sido, cuando menos, intrigantes. Una lección importante de la aparición del lenguaje artificial es que no debemos asumir que la forma en que los humanos resolvemos las tareas comunicativas es la única solución posible. Encontramos un ejemplo conocido en Bouchacourt y Baroni (2018), quienes utilizaron una configuración similar a la de la Figura 1. Después de que los agentes crearan su propio sistema comunicativo, Bouchacourt y Baroni pusieron a prueba con qué pericia podían hablar sobre las imágenes utilizadas en su entrenamiento (perros, aviones y alimentos como los que se muestran en la Figura 1) o bien hablar sobre versiones distorsionadas de esas imágenes que no habían visto anteriormente. Las distorsiones parecían, básicamente, estática de televisión de colores. Sorprendentemente, los agentes artificiales podían comunicarse casi igual de bien sobre las imágenes distorsionadas que sobre las originales. ¿Cómo es posible que un lenguaje artificial creado para hablar de imágenes de perros, comida y aviones sirva también para hablar de versiones

distorsionadas que no se parecen en nada al original? Este desconcertante resultado se explica porque estos entes no piensan como las personas. Probablemente, los seres humanos crearían un lenguaje que les permitiera hablar de los objetos mostrados en las imágenes: *perros*, *bagels*, etc. Los agentes artificiales, en cambio, parecían haber creado un lenguaje con el que podían hablar de características visuales superficiales de las imágenes. Es decir, hablaban de cosas como la tonalidad de determinados píxeles de las imágenes, en lugar de sobre si mostraban bagels, perros o aviones. Estas características superficiales seguían presentes en las imágenes distorsionadas, pero los humanos somos incapaces de percibir las. La lección es que, mientras que los seres humanos interpretamos y describimos imágenes de forma natural basándonos en los objetos o escenas que representan, a los agentes artificiales les basta con hablar de píxeles y colores. Lo que a nosotros y a ellos nos resulta natural puede ser bien diferente. Por lo tanto, si queremos que inventen lenguajes similares a los nuestros, tenemos que ingeniárnoslas para que vean el mundo como lo vemos nosotros.

Otro hallazgo en esta línea es que las lenguas artificiales creadas de este modo tampoco tienen la tendencia de las lenguas humanas a la concisión. Una huella estadística universal común a todas las lenguas humanas es que las palabras más frecuentes son las más cortas. Por ejemplo, en inglés los artículos *the* y *a* son palabras muy frecuentes y cortas en comparación con *romboidal*, un término bas-

tante infrecuente. Esto es eficaz porque ahorra esfuerzo a hablantes y oyentes: las palabras que más utilizamos son las más cortas y fáciles de emitir. Los agentes artificiales tienen la tendencia opuesta (Chaabouni et al., 2019). Tienen a expresar conceptos frecuentes con las palabras más largas posibles y, a su vez, todas las palabras tienden a ser más largas de lo necesario. La razón de este comportamiento es sencillamente que, mientras a los humanos nos importa que las palabras o las frases sean muy largas, a una máquina no. Por el contrario, las cadenas más largas les permiten crear mensajes más fáciles de interpretar por otros agentes artificiales, porque les proporcionan más caracteres con los que codificar el concepto pretendido.

Los anteriores son solo dos ejemplos entre una gran cantidad de descubrimientos de esta disciplina incipiente. Al principio, puede parecer obvio que, para que las lenguas artificiales se parezcan más a los humanos, sus usuarios deben preocuparse por lo mismo que nosotros (en los ejemplos anteriores: *reducir la longitud de las palabras y hablar de objetos y no de píxeles*). Sin embargo, antes de realizar estos estudios, es mucho menos evidente qué es exactamente lo que importa de cómo interactuamos y procesamos la información. En otras palabras, estos resultados nos enseñan valiosas lecciones sobre qué partes de la experiencia humana dan forma al lenguaje. Esto las convierte en un campo de pruebas ideal para estudiar la evolución del lenguaje.

■ EL FUTURO DE APRENDER DEL PASADO

Este artículo aborda algunas formas novedosas de estudiar la evolución del lenguaje: a través de datos a gran escala y mediante experimentos sobre la creación de lenguaje artificial. Ambas son posibles gracias a recursos y métodos informáticos de los que no disponíamos hasta hace pocos años. No obstante, los viejos retos continúan vigentes. El estudio de la evolución del lenguaje sigue siendo abductivo y basado en hipótesis, y es poco probable que logremos aislar los ingredientes y procesos concretos que la constituyen. Al fin y al cabo, el lenguaje humano es el producto complejo de muchos factores relacionados. En cualquier caso, la caja de herramientas de que disponemos está en constante expansión y nos permite seguir perfeccionando y cuestionando las teorías actuales.

Al igual que ocurre con el estudio del pasado, es difícil predecir la evolución futura de este campo. De lo que podemos estar relativamente seguros es de que la diversidad lingüística desempeñará un papel cada vez más importante en los próximos años. Los enfoques computacionales actuales suelen requerir una cantidad ingente de datos. Por ello, la mayor parte de la investigación actual a gran escala se basa en unas pocas lenguas seleccionadas con una extensa representación en registros tanto escritos

como orales. Sin embargo, los esfuerzos por ofrecer una imagen menos sesgada y a escala de la diversidad lingüística del mundo –como el corpus DoReCo (Seifart et al., 2022) o BLOOM (BigScience Workshop, 2023)– indican que las cosas están cambiando también en este sentido. La continua digitalización de lenguas y dialectos infra-representados probablemente aporte una gran cantidad de datos nuevos que nos permitirán poner a prueba ideas antiguas y nuevas. Este campo sigue evolucionando, al igual que las lenguas.

REFERENCIAS

- Bouchacourt, D., & Baroni, M. (2018). How agents see things: On visual representations in an emergent language game. En E. Riloff, D. Chiang, J. Hockenmaier & J. Tsujii, *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing* (pp. 981–985). Association for Computational Linguistics.
- BigScience Workshop. (2023). BLOOM: A 176B-parameter open-access multilingual language model. arXiv. <https://doi.org/10.48550/arxiv.2211.05100>
- Brochhagen, T., & Boleda, G. (2022). When do languages use the same word for different meanings? The Goldilocks principle in colexification. *Cognition*, 226, 105179. <https://doi.org/10.1016/j.cognition.2022.105179>
- Brochhagen, T., Boleda, G., Gualdoni, E., & Xu, Y. (2023). From language development to language evolution: A unified view of human lexical creativity. *Science*, 381(6656), 431–436. <https://doi.org/10.1126/science.ade7981>
- Chaabouni, R., Kharitonov, E., Dupoux, E., & Baroni, M. (2019). Anti-efficient encoding in emergent communication. En *Proceedings of NeurIPS 2019 (33rd Conference on Neural Information Processing Systems)* (pp. 6290–6300). Curran Associates.
- Corballis, M. C. (2008). Not the last word. *American Scientist*, 96(1), 68–70.
- Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., & Fei-Fei, L. (2009). ImageNet: A large-scale hierarchical image database. En *IEEE Computer Vision and Pattern Recognition (CVPR)* (pp. 248–255). <https://doi.org/10.1109/CVPR.2009.5206848>
- Kemp, C., & Regier, T. (2012). Kinship categories across languages reflect general communicative principles. *Science*, 336(6084), 1049–1054. <https://doi.org/10.1126/science.1218811>
- Lazaridou, A., & Baroni, M. (2020). Emergent multi-agent communication in the deep learning era. arXiv. <https://doi.org/10.48550/arxiv.2006.02419>
- Rzyski, C., Tresoldi, T., Greenhill, S. J., Wu, M.-S., Schweikhard, N. E., Kop-tjevskaja-Tamm, M., Gast, V., Bodt, T. A., Hantgan, A., Kaiping, G. A., Chang, S., Lai, Y., Morozova, N., Arjava, H., Hübner, N., Koile, E., Pepper, S., Proos, M., Van Epps, B., ... List, J.-M. (2020). The database of cross-linguistic colexifications, reproducible analysis of cross-linguistic polysemies. *Scientific Data*, 7, 13. <https://doi.org/10.1038/s41597-019-0341-x>
- Seifart, F., Paschen, L., & Stave, M. (2022). Language Documentation Reference Corpus (DoReCo) 1.2. Berlin & Lyon: Leibniz-Zentrum Allgemeine Sprachwissenschaft & laboratoire Dynamique Du Langage (UMR5596, CNRS & Université Lyon 2). <https://doi.org/10.34847/nkl.7cbfq779>
- Passmore, S., Barth, W., Greenhill, S. J., Quinn, K., Sheard, C., Argyriou, P., Birchall, J., Bowern, C., Calladine, J., Deb, A., Diederer, A., Metsäranta, N. P., Araujo, L. H., Schembri, R., Hickey-Hall, J., Honkola, T., Mitchell, A., Poole, L., Rácz, P. M., ... Jordan, F. M. (2023). Kinbank: A global database of kinship terminology. *PLOS ONE*, 18(5), e0283218. <https://doi.org/10.1371/journal.pone.0283218>
- Xu, Y., Duong, K., Malt, B. C., Jiang, S., & Srinivasan, M. (2020). Conceptual relations predict colexification across languages. *Cognition*, 201, 104280. <https://doi.org/10.1016/j.cognition.2020.104280>
- Zaslavsky, N., Kemp, C., Regier, T., & Tishby, N. (2018). Efficient compression in color naming and its evolution. *Proceedings of the National Academy of Sciences*, 115(31), 7937–7942. <https://doi.org/10.1073/pnas.1800521115>

THOMAS BROCHHAGEN. Profesor ayudante doctor en Ciencia Cognitiva Computacional en el Departamento de Traducción y Ciencias del Lenguaje de la Universitat Pompeu Fabra (España). Sus áreas de investigación incluyen la evolución del lenguaje, la inteligencia artificial, los modelos bayesianos y la estadística. ✉ thomas.brochhagen@upf.edu