

Curso- taller:

Fundamentos de la comparación histórica y tipológica del lenguaje: enfoques cuantitativos y cualitativos

A cargo del Dr. Johann-Mattis List.

Max Planck Institute for the Science of Human History (Jena, Alemania).

La investigación en lingüística histórica y tipológica ha cambiado drásticamente en los últimos años. En Alemania, por ejemplo, muchos institutos de lingüística indoeuropea de enfoque tradicional han desaparecido, pero al mismo tiempo, han surgido nuevos equipos de investigación para los que los métodos cuantitativos juegan un rol fundamental.

No obstante, a pesar de la promesa de estas investigaciones recientes, los métodos cuantitativos aún no han logrado producir por sí mismos conocimientos novedosos sobre la historia de las lenguas del mundo. La mayoría de ellos están todavía en ciernes y hasta el momento solo permiten resolver una pequeña fracción de los problemas clásicos de la lingüística histórica y tipológica. Por ejemplo, las computadoras no pueden sustituir a los lingüistas en el descubrimiento de cognados entre idiomas diferentes. Tampoco pueden realizar una reconstrucción lingüística sin ayuda humana. Otro problema es que los métodos computacionales son falibles y cometen errores que los lingüistas bien formados nunca los cometerían. Por último, la construcción de base de datos con fines comparativos no puede realizarse sólo mediante métodos automáticos sino que requiere del trabajo de lingüistas expertos que seleccionen e interpreten los datos disponibles en las gramáticas existentes.

En síntesis, mientras que los enfoques meramente computacionales y cuantitativos son comunes hoy en día, el presente curso se enfoca en la comunicación entre la lingüística clásica y la computacional mediante el desarrollo de métodos que permitan a los lingüistas producir sus datos en formatos de computadora y al mismo tiempo presentar los resultados en formatos humanamente legibles.

El curso se propone brindar a los estudiantes en lingüística las herramientas y conceptos básicos de los nuevos métodos cuantitativos que permiten complementar el enfoque cualitativo en lingüística histórica y tipológica. En ese contexto, la atención se va a centrar en temas como la creación de diccionarios etimológicos (cf. List, Greenhill, and Gray 2017, List et al. 2018), la comparación de secuencias en lingüística histórica (List, 2014); List 2018d), la reconstrucción filogenética (Greenhill 2015), la tipología léxica (List et al. 2018), la estandarización (Forkel et al. 2018), la curación de datos (List, Greenhill, and Gray 2017, Hill and List 2017), y la importancia de la estadística y de los procesos estocásticos en la comparación de las lenguas del mundo (Kessler 2001).

El curso consiste en tres partes. En la primera se presentará la teoría de los nuevos métodos cuantitativos y se los comparará con los métodos tradicionales. La segunda parte será de carácter práctico y se enfocará en las diferentes posibilidades de uso de los nuevos métodos cuantitativos. Por último, la tercera estará destinada a la presentación y discusión de los problemas específicos de las investigaciones llevadas a cabo por los estudiantes.

Se espera que los doctorandos puedan llegar a evaluar críticamente los nuevos soportes computacionales (como la reconstrucción filogenética o las pruebas de significación) a la vez que incorporen un conocimiento adecuado de los métodos lingüísticos clásicos.

Así también, mediante la resolución de ejercicios específicos, se instruirá a los estudiantes a aplicar los enfoques cuantitativos y cualitativos de una forma independiente.

Por último, el curso incluye una breve introducción a la programación para lingüistas, con énfasis en Python (un lenguaje de programación de propósito general preparado para realizar cualquier tipo de programa, incluso, páginas web).

Sobre Johann-Mattis List

Dr. Johann-Mattis List es científico principal del Departamento de Lingüística y Evolución Cultural del Instituto Max Planck para la Ciencia de la Historia Humana. Allí lleva a cabo un proyecto de investigación interdisciplinario sobre la Comparación de Lenguas con Asistencia de Computadoras (Computer-Assisted Language Comparison, CALC), financiado por el Consejo Europeo de Investigación (European Research Council, ERC, 04/2017-03/2022). En este proyecto, desarrolla métodos e interfaces que cierran la brecha entre los enfoques computacional y clásico de la lingüística histórica. En su investigación, el Dr. List generalmente sigue una perspectiva cuantitativa, empírica y basada en datos sobre el cambio de lenguaje. Sin embargo, en contraste con los enfoques puramente computacionales, intenta mantener su investigación estrechamente conectada con teorías clásicas y tradicionales. De esta manera, el marco de sus investigaciones es con asistencia de computadoras, y no basado en computadoras.

Literatura mencionada y recomendada para introducción

Forkel, Robert, Johann-Mattis List, Simon J. Greenhill, Christoph Rzymiski, Sebastian Bank, Michael Cysouw, Harald Hammarström, Martin Haspelmath, Gereon A. Kaiping, and Russell D. Gray. 2018. “Cross-Linguistic Data Formats, Advancing Data Sharing and Re-Use in Comparative Linguistics.” *Scientific Data* 5 (180205): 1–10.

Greenhill, Simon. 2015. “Evolution and Language: Phylogenetic Analyses.” In *International Encyclopedia of the Social & Behavioral Sciences (Second Edition)*, edited by James D. Wright, Second Edition, 370–77. Oxford: Elsevier. doi:<https://doi.org/10.1016/B978-0-08-097086-8.81035-1>.

Hill, Nathan W., and Johann-Mattis List. 2017. “Challenges of Annotation and Analysis in Computer-Assisted Language Comparison: A Case Study on Burmish Languages.” *Yearbook of the Poznań Linguistic Meeting* 3 (1): 47–76.

Kessler, Brett. 2001. *The Significance of Word Lists: Statistical Tests for Investigating Historical Connections Between Languages*. Stanford: CSLI Publications.

List, Johann-Mattis. 2014. *Sequence Comparison in Historical Linguistics*. Düsseldorf: Düsseldorf University Press.

List, Johann-Mattis, Simon J. Greenhill, and Russell D. Gray. 2017. “The Potential of Automatic Word Comparison for Historical Linguistics.” *PLOS ONE* 12 (1). Public Library of Science: 1–18.

List, Johann-Mattis, Simon J. Greenhill, Cormac Anderson, Thomas Mayer, Tiago Tresoldi, and Robert Forkel. 2018. “CLICS². An Improved Database of Cross-Linguistic Colexifications Assembling Lexical Data with Help of Cross-Linguistic Data Formats.” *Linguistic Typology* 22 (2): 277–306.