

## The Limits of the Comparative Method: Innovative Approaches to Understanding Orphan Languages

57<sup>th</sup> Annual Meeting of the *Societas Linguistica Europaea*  
21 – 24 August 2024  
University of Helsinki

Abbie Hantgan: Langage, Langues et Cultures d’Afrique (LLACAN), CNRS [abbie.hantgan-sonko@cirs.fr](mailto:abbie.hantgan-sonko@cirs.fr)  
Rik van Gijn: Center for Linguistics, Leiden University [e.van.gijn@hum.leidenuniv.nl](mailto:e.van.gijn@hum.leidenuniv.nl)

### 1. Description

This workshop aims to explore and challenge the Comparative Method's limitations in historical-comparative linguistics, focusing on the study of 'orphan languages'. The Comparative Method (CM) stands as one of the most impactful methodologies in historical-comparative linguistics, revolutionizing our understanding of the development of language families. Its enduring relevance in modern linguistics attests to its efficacy. However, the CM comes with a significant limitation: it requires *comparanda*—distinct languages suspected to share a lineage. This constraint renders it inapplicable for studying language isolates and many small language families, which, in reality, make up the majority of language families. We refer to these linguistically unique or small-scale genealogical units as 'orphan languages'.

While we join others in recognizing the importance of these 'orphan languages' to historical linguistics, they are often conspicuously absent from discussions that critique the limitations, or offer alternatives, to the Comparative Method. That is, the focus in terms of remnants of linguistic diversity remains on single languages classified as 'isolates'—those with no known genealogical relatives. Language isolates in particular have been extensively cataloged and analyzed from various angles—including comparative, typological, and diachronic approaches (Georg 2015, Epps & Michael 2023, Salaberry et al. 2020, Campbell 2017)—as well as from geographic perspectives (Urban 2021). Emerging research is adopting cross-disciplinary methods, utilizing statistical patterning (Van Gijn et al. 2022) and Natural Language Processing (NLP) in corpus linguistics (Gamallo et al. 2020). Recent investigations actively merge linguistics with genetics (Matsumae et al. 2021, Laurent et al. 2023) to tackle the tough questions of the origins the speakers of language *isolates*, but not necessarily the broader category of language *orphans*.

Furthermore, despite advancements, the limitations of the (classical) CM are still evident: it cannot effectively trace the history of language orphans and their speakers, and there is no alternative method with proven results that can take its place. Additionally, many potential language orphans may exist in under-explored regions, or some languages might be wrongly assumed to have affiliations with their neighbors. According to Glottolog (Hammarström et al. 2023), there are at least 200 language groups considered to be 'top-level families' with fewer than ten languages each, and a full 121 languages that are simply labeled as 'unclassifiable'. This striking statistic highlights a significant gap in our understanding of linguistic diversity. With these languages remaining beyond the reach of historical linguistic methods, achieving a comprehensive classification of the world's languages remains an elusive goal. Our workshop aims to confront this challenge head-on.

The *Orphan languages* workshop is designed with two key objectives in mind. First, we aim to broaden the academic discourse beyond the current state of the art, which predominantly focuses on language isolates. Our goal is to introduce and define the concept of 'orphan languages,' a term that encompasses not only language isolates but also clusters of languages whose classifications are

unknown or disputed. By expanding this definition, we acknowledge a wider spectrum of linguistic enigmas and affirm the importance of studying these languages to understand human linguistic heritage more fully. Second, the *Orphan Languages* workshop aims to address the above highlighted challenges by exploring innovative methods for reconstructing the proto-stages of language isolates and unaffiliated language groups, circumventing the limitations of the classical comparative method. Addressing an underexplored area in linguistic research, this workshop proposes to bring together experts to discuss innovative methodologies for studying languages with unclear genealogical relationships.

The contributions to this workshop address a wide range of topics and approaches to studying orphan languages: Some contributions take a quantitative, large-scale approach (Van Gijn & Norder, Blum & List), while others focus on highly detailed, in-depth studies (Van Schie, Broadwell, Coburn & Sands, Dyachkov). Some approaches focus on the geographical context of orphan languages (Van Gijn & Norder, Van Schie, Zamponi, Badagil, Hantgan), while others emphasize the genealogical context (Blum & List, Broadwell, Dyachkov, Martínez-Areta). A number of contributions explicitly advocate multi-disciplinary approaches to reconstruct the histories of orphan languages (Van Gijn & Norder, Bardagil, Broadwell, Zuk, Hantgan), others focus particularly on how new approaches to linguistic structures reveal histories of orphan languages (Van Schie, Coburn & Sands, Zamponi, Martínez-Areta, Dyachkov). Finally, some studies focus on the extent to which orphan languages may offer us an insight into older layers of areal history (Van Gijn & Norder, Bardagil, Hantgan).

This interdisciplinary workshop constitutes a confluence of scholarly work that spans the global expanse of linguistic diversity, encompassing research on African, European, Southeast Asian, North American, and South American languages. The contributors, who are leading figures in their respective fields, will interrogate complex strategies of language evolution, contact, and divergence in varied sociocultural and geographical contexts. The workshop's scope, embracing both area-specific and cross-linguistic studies, is poised to contribute substantively to the epistemological expansion of historical linguistics.

With a rigorous application of interdisciplinary methodologies, the workshop is designed to further the academic community's knowledge of orphan languages. It will scrutinize the current limitations of the Comparative Method, advocate for more inclusive linguistic categorizations, and explore innovative strategies for historical reconstruction. The goal is to refine the analytical tools that linguists employ to trace the evolution of languages that lack extensive written records or clear genealogical ties.

In essence, this workshop is an endeavor to systematically integrate orphan languages into the domain of historical linguistics. It aspires to stimulate a scholarly reassessment of the frameworks used to understand linguistic development and diversity, ensuring that orphan languages are not merely footnotes in the broader narrative of language history but are recognized as key to unlocking patterns of human communication and migration.

#### Keywords

Comparative Linguistics, Orphan Languages, Language Isolates, Historical Linguistics, Interdisciplinary Methods

## References

- Campbell, Lyle. (ed.). 2017. *Language Isolates*. London and New York: Routledge.
- Epps, Patience & Lev, Michael (eds.). 2023. *Language Isolates I: Aikanã to Kandozi-Shapra: An International Handbook (Amazonian Languages)*. Vol. 1. 2 vols. Berlin, Boston: De Gruyter Mouton. (<https://www.degruyter.com/serial/hskal-b/html>)
- Gamallo, Pablo & Pichel, José Ramon & Alegria, Iñaki. 2020. Measuring Language Distance of Isolated European Languages. *Information* 11(4). (doi:10.3390/info11040181) (<https://www.mdpi.com/2078-2489/11/4/181>)
- Georg, Stefan (ed.). 2015. *Language isolates* (4 volumes). London, England: Routledge.
- Hammarström, Harald & Forkel, Robert & Haspelmath, Martin & Bank, Sebastian. 2023. *Glottolog* 4.8. Leipzig: Max Planck Institute for Evolutionary Anthropology. <https://doi.org/10.5281/zenodo.8131084> (Available online at <http://glottolog.org>, Accessed on 2023-11-01.)
- Laurent, Romain & Szpiech, Zachary A & da Costa, Sergio S & Thouzeau, Valentin & Fortes-Lima, Cesar A & Dessarps-Freichy, Françoise & Lémée, Laure et al. 2023. A genetic and linguistic analysis of the admixture histories of the islands of Cabo Verde. *eLife. eLife Sciences Publications, Ltd* 12. e79827. (Ed. Huerta-Sanchez, Emilia & Przeworski, Molly & Huerta-Sanchez, Emilia & Zhang, Xinjun.) (doi:10.7554/eLife.79827)
- Matsumae, Hiromi & Ranacher, Peter & Savage, Patrick E. & Blasi, Damián E. & Currie, Thomas E. & Koganebuchi, Kae & Nishida, Nao et al. 2021. Exploring correlations in genetic and cultural variation across language families in northeast Asia. *Science Advances*. American Association for the Advancement of Science 7(34). eabd9223. (doi:10.1126/sciadv.abd9223)
- Salaberri, Iker & Krajewska, Dorota & Reguero, Urtzi & Zuloaga, Eneko & Duhalde, Maitena (collaborator) & Uribe-Etxebarria, Oxel (collaborator) & Monforte, Sergio (collaborator) & Santazilia, Ekaitz (collaborator) & Ariztimuño, Borja (collaborator) & Ulibarri, Koldo (collaborator). 2020. *Investigating language isolates: typological and diachronic perspectives*. Zenodo. (doi:10.5281/zenodo.3985030)
- Urban, Matthias. 2021. The geography and development of language isolates. *Royal Society Open Science* 8(202232). (doi:10.1098/rsos.202232)
- Van Gijn, Rik & Norder, Sietze & Arias, Leonardo & Emlen, Nicholas Q. & Azevedo, Matheus C. B. C. & Caine, Allison & Dunn, Saskia et al. 2022. The social lives of isolates (and small language families): the case of the Northwest Amazon. *Interface Focus. Royal Society* 13(1). 20220054. (doi:10.1098/rsfs.2022.0054)