

# Phonological databases

Convener: Ian Joo

*Nagoya University of Commerce and Business*

October 28, 2023

Phonological typology in the 21<sup>st</sup> century is endowed with an increasing number of phonological databases, containing various types of phonological information, such as phonemic inventories (Maddieson 2009; Moran and McCloy 2019), phonotactic constraints (Mielke 2008; Maddieson, Flavier, et al. 2013; Nikolaev 2018; Joo and Hsu 2023), loan segments (Grossman et al. 2020), and ancient sounds (Moran, Grossman, et al. 2021). These databases open door for a myriad of data-driven approaches on phonological universals and areal patterns. Phonological typologists may benefit from a methodological discussion on how to profit from the possibilities of such large data.

To host such a discussion in a tentative workshop at the 57<sup>th</sup> annual meeting of the *Societas Linguistica Europaea*, abstracts related to the following questions and other related topics are called for:

- **How to use the databases?** The phonological databases can be used for diverse purposes, such as detecting the areal distribution of consonant inventories (Nikolaev 2019) or investigating the correlation between climate and phonology (Maddieson and Benedict 2023). With our creativity, the available databases can be exploited yet further for endless research possibilities, especially when we combine the phonological databases with the non-phonological ones, such as Grambank (Skirgård et al. 2023).
- **How can the databases complement each other?** Each database, by the nature of its design and the first-hand data it relies on, is limited in one way or another. Anderson et al. (2023) have shown that the descriptions of the same sample of lects are alarmingly inconsistent across different databases. Employing multiple databases, therefore, can be the desired methodology in order to prevent a database becoming a data bias. How, then, can we extract cross-compatible data from differently designed databases, with different lect classifications, phonological transcriptions, and theoretical backgrounds?
- **What's next?** What are some phonological (or phonetic) properties of human language that are not covered by the databases published so far? For example, no currently available database contains extensive data on prosody, vowel formant frequencies, sign language phonology, morphophonological rules, and many other important phonetic-phonological variables. If you are working on a database that fills in one of these gaps, then this venue can be the right place to present your ongoing research and receive feedback from other linguists working on previous databases.

Please send your abstract (up to 300 words excluding references, 12pt Times New Roman, 2.5cm margin, pdf format) on one of these or other related topics to [ian\\_joo@nucba.ac.jp](mailto:ian_joo@nucba.ac.jp) by **19 November 2023**. If I receive eight or more abstracts, I will send a workshop proposal to the Chair of SLE workshops for their approval.

## References

- Anderson, Cormac, Tiago Tresoldi, Simon J. Greenhill, Robert Forkel, Russell D Gray, and Johann-Mattis List (2023). “Measuring variation in phoneme inventories”. In: *Journal of Language Evolution* (accepted).
- Grossman, Eitan, Elad Eisen, Dmitry Nikolaev, and Steven Moran (2020). “SegBo: A database of borrowed Sounds in the world’s languages”. In: *Proceedings of the 12th language resources and evaluation conference*. European Language Resources Association, pp. 5316–5322.
- Joo, Ian and Yu-Yin Hsu (2023). “Phonotacticon: a cross-linguistic phonotactic database”. In: *Language Resources and Evaluation* (under review). DOI: 10.21203/rs.3.rs-3269302/v1.
- Maddieson, Ian (2009). *Patterns of sounds*. Cambridge University Press.
- Maddieson, Ian and Karl Benedict (2023). “Demonstrating environmental impacts on the sound structure of languages: challenges and solutions”. In: *Frontiers in Psychology* 14.
- Maddieson, Ian, Sébastien Flavier, Egidio Marsico, Christophe Coupé, and François Pellegrino (2013). “LAPSyd: Lyon-Albuquerque phonological systems database”. In: *Interspeech 2013*. International Speech Communication Association (ISCA). DOI: 10.21437/interspeech.2013-660.
- Mielke, Jeff (2008). *The emergence of distinctive features*. Oxford University Press.
- Moran, Steven, Eitan Grossman, and Annemarie Verkerk (2021). “Investigating diachronic trends in phonological inventories using BDPROTO”. In: *Language Resources and Evaluation* 55.1, pp. 79–103.
- Moran, Steven and Daniel McCloy (2019). *PHOIBLE 2.0*. Max Planck Institute for the Science of Human History. URL: <https://phoible.org/>.
- Nikolaev, Dmitry (2018). “The Database of Eurasian Phonological Inventories: A research tool for distributional phonological typology”. In: *Linguistics Vanguard* 4.1.
- (2019). “Areal dependency of consonant inventories”. In: *Language Dynamics and Change* 9.1, pp. 104–126.
- Skirgård, Hedvig et al. (2023). “Grambank reveals the importance of genealogical constraints on linguistic diversity and highlights the impact of language loss”. In: *Science Advances* 9.16, eadg6175. DOI: 10.1126/sciadv.adg6175.