

Introduction to lectometry

Ian Joo

Otaru University of Commerce

May 20, 2026

This chapter introduces the reader to LECTOMETRY, the quantifiable degree of dissimilarity between LECTS. A lect is defined as any level of variety of the human language, including what are commonly referred to as “dialects” or “languages”, as well as many other levels, such as sociolects, idiolects, situlects, doculects, and so on. Dissimilarity can be defined in many ways, namely in the domains of phonology, semantics, syntax, vocabulary, genealogy, register, and so on.

The term “lectometry” was coined by Ruetter (2012) and further used by a number of papers (Ruetter et al. 2011; Geeraerts 2018; Plevoets 2020, e.g.), defined as “measur[ing] distances between lects, where lects are defined as language varieties that differ along any of various dimensions, such as dialects and regiolects” (Rosseel et al. 2020, p. 132). However, the first authors tended to distinguish “lects” from “languages” and reserve the term “lectometry” to distance between dialectal or sub-dialectal varieties.

Here we expand the term (without contradicting the original definition) by interpreting “lects” as including not only dialects, regiolects, and sociolects but literally any level of variation, namely those that are commonly referred to as “languages” or even “macrolanguages”. We highlight that this is not a redefinition of the term but only a logical corollary to the fact that there is no consistently defined boundary between what is called a “dialect” and a “language”, as this dichotomy is a modern European social construct (Van Rooy 2020), and that even the difference between unrelated “languages” can be seen as a variation of the human language (or the human linguistic capacity).

The novelty of this monograph lies not in merely expanding the scope of the term lectometry, but rather **finding a common ground between the two fields of dialectology and quantitative typology**. Dialectology commonly employs various methods of DIALECTOMETRY (Nerbonne and W. Heeringa 1997; W. J. Heeringa 2004, e.g.). Quantitative typology, on the other hand, also commonly involves comparing different “languages” at a numerical scale, which often results in quantified levels of difference, i.e. distance (Szeto and Yurayong 2021; Joo and Hsu ahead of print, e.g.). Logically, if there is no consistent boundary between “languages” and “dialects”, there should also not be a strict barrier between dialectometry of dialectology and “language-metry” of quantitative typology (although, of course, different metrics may be used depending on how genealogically close two varieties are). Thus there is the need to relabel the two fields as one and revisit it within a common frame.

The chapter concludes by summarizing the following chapters, which will be done after their first drafts are written.

References

- Geeraerts, Dirk (2018). “A lectometric definition of lexical destandardization”. In: *Wortschätze: Dynamik, Muster, Komplexität*. Ed. by Stefan Engelberg et al. Berlin: De Gruyter, pp. 233–244. DOI: 10.1515/9783110579963-013.
- Heeringa, Wilbert Jan (2004). “Measuring dialect pronunciation differences using Levenshtein distance”. PhD thesis. University of Groningen.
- Joo, Ian and Yu-Yin Hsu (ahead of print). “Phonological distances between Eurasian lects measured via Phonotacticon 1.0 reveal areal patterns”. In: *Linguistics*. DOI: 10.1515/ling-2024-0157.
- Nerbonne, John and Wilbert Heeringa (1997). “Measuring dialect distance phonetically”. In: pp. 11–18.
- Plevoets, Koen (2020). “Lectometry and Latent Variables: a Model for Underlying Determinants of (Normative) Choices in Written and Audiovisual Translations”. In: *Zeitschrift für Dialektologie und Linguistik* 87.2, pp. 144–172. DOI: 10.25162/zdl-2020-0006.
- Rosseel, Laura, Karlien Franco, and Melanie Röthlisberger (2020). “Extending the Scope of Lectometry”. In: *Zeitschrift für Dialektologie und Linguistik* 87.2, pp. 131–143. DOI: 10.25162/zdl-2020-0005.
- Ruette, Tom (2012). “Aggregating lexical variation: towards large-scale lexical lectometry”. PhD thesis. University of Leuven.
- Ruette, Tom, Dirk Speelman, and Dirk Geeraerts (2011). “Measuring the lexical distance between registers in national varieties of Dutch”. In: *Línguas pluricêntricas : Variação linguística e dimensões sociocognitivas = Pluricentric languages : Linguistic variation and sociocognitive dimensions*. Ed. by Augusto Soares da Silva, Amadeu Torres, and Miguel Gonçalves. Braga: Faculdade de Filosofia, Universidade Católica Portuguesa, pp. 541–554.
- Szeto, Pui Yiu and Chingduang Yurayong (2021). “Sinitic as a typological sandwich: Revisiting the notions of Altaicization and Taicization”. In: *Linguistic Typology* 25.3, pp. 551–599. DOI: 10.1515/lingty-2021-2074.
- Van Rooy, Raf (2020). *Language or dialect?: The history of a conceptual pair*. Oxford: Oxford University Press.