

# Dmitry Nikolaev, curriculum vitae

Visiting researcher  
Department of Linguistics, Stockholm University  
SE - 10691 Stockholm, Sweden

dmitry.nikolaev@ling.su.se  
orcid.org/0000-0002-3034-9794

<https://scholar.google.com/citations?user=My18EpkAAAAJ&hl=en>  
<https://su-se.academia.edu/DmitryNikolaev>  
<https://github.com/macleginn>  
<https://dnikolaev.com>

## Education

- 2012 Ph.D. Philology  
Russian State University for the Humanities
- 2008 Specialist (BA+MA) degree in Philology, Russian State University for the Humanities

## Academic positions

- 2019–2021 **Stockholm University**  
Visiting researcher, computational linguistics group, Department of Linguistics
- 2018–2019 **Hebrew University of Jerusalem**  
Postdoc, NLP lab, Department of Computer Science / Department of Linguistics
- 2016–2017 **Hebrew University of Jerusalem**  
Postdoc and part-time teaching fellow, Department of Linguistics
- 2012–2013 **Lomonosov Moscow State University**  
Senior researcher, Department of Germanic and Celtic Philology, Philological Faculty
- 2012–2016 **Russian State University for the Humanities**  
Researcher, Centre for Typology and Semiotics of Folklore

## Visiting scientist

- 2015 Department of Linguistics, MPI Leipzig

## Publications

### Peer-reviewed journal articles and conference papers

- Accepted** Koptjevskaja-Tamm, Maria and Dmitry Nikolaev. Talking about temperature and social thermoregulation in the languages of the world. *International Review of Social Psychology*.

- Accepted** Georgakopoulos, Thanasis, Eitan Grossman, Dmitry Nikolaev and Stéphane Polis. Universal and macro-areal patterns in the lexicon: A case-study in the perception-cognition domain. *Linguistic Typology*.
- Accepted** Nikolaev, Dmitry and Mikhail Shumilin. Some considerations on the attribution of the “New Apuleius”. *Classical Quarterly*.
- Accepted** Nikolaev, Dmitry and Mikhail Shumilin. Identifying Latin authors through maximum-likelihood Dirichlet inference: A contribution to model-based stylometry. *Shagi/Steps*.
- 2020 Nikolaev, Dmitry and Eitan Grossman. Consonant co-occurrence classes and the feature-economy principle. *Phonology* 37(3), 419–451.
- 2020 Nikolaev, Dmitry. On the sound change \**ld* > *nd* in Tibetic: A reply to Xun Gong. *Cahiers de Linguistique Asie Orientale* 49(2), 250–255.
- 2020 Nikolaev, Dmitry, Ofir Arviv, Taelin Karidi, Neta Kenneth, Veronika Mitnik, Lilja Maria Saeboe and Omri Abend. Fine-grained analysis of cross-linguistic syntactic divergences. *Proceedings of ACL*, 1159–1176.
- 2020 Nikolaev, Dmitry, Taelin Karidi, Neta Kenneth, Veronika Mitnik, Lilja Maria Saeboe and Omri Abend. Morphosyntactic predictability of translationese. *Linguistics Vanguard* 6(1), 20190077.
- 2020 Grossman, Eitan, Elad Eisen, Dmitry Nikolaev and Steven Moran. SegBo: A database of borrowed sounds in the world’s languages. *Proceedings of LREC*, 5318–5324.
- 2020 Choshen, Leshem\*, Dmitry Nikolaev\*, Yevgeni Berzak and Omri Abend. Classifying syntactic errors in learner language. *Proceedings of CoNLL*, 97–107. \*Shared first authorship.
- 2020 Nikolaev, Dmitry. A review of Jenseit G.B., McGillivray B. *Quantitative historical linguistics: A corpus framework*. Oxford University Press, 2017. *Voprosy Jazykoznanija* 1, 155–160.
- 2019 Nikolaev, Dmitry. Areal dependency of consonant inventories. *Language Dynamics and Change* 9(1), 104–126.
- 2018 Nikolaev, Dmitry. The Database of Eurasian Phonological Inventories: A research tool for distributional phonological typology. *Linguistics Vanguard* 4(1), 20170050.
- 2018 Nikolaev, Dmitry and Eitan Grossman. Areal sound change and the distributional typology of affricate richness in Eurasia. *Studies in Language* 42(3), 562–599.
- 2018 Nikolaev, Dmitry. A review of Hickey, Raymond (ed.). *The Cambridge handbook of areal linguistics*. Cambridge University Press, 2017. *Voprosy Jazykoznanija* 1, 126–129.
- 2016 Nikolaev, Dmitry. A new algorithm for extracting formulas from poetic texts and the formulaic density of Russian *bylinas*. *Oral Tradition* 30(1), 111–136.
- 2013 Nikolayev, Dmitry. A review of Stüber, Karin; Zehnder, Thomas; Remmer, Ulla: Indogermanische Frauennamen. Heidelberg: Winter, 2009. *Zeitschrift für celtische Philologie* 60(1), 335–337.

## Papers in edited volumes

- Accepted** Harel, Halely, Orly Goldwasser and Dmitry Nikolaev. Mapping the ancient mind: Introducing iClassifier, a new platform for systematic analysis of classifiers in Egyptian and beyond. In: eds. Joshua Roberson, Steve Vinson, and Rita Lucarelli. *Proceedings of Ancient Egypt and New Technology: The Present and Future of Computer Visualization, Virtual Reality and other Digital Humanities in Egyptology*. 29–30 March 2019 at Indiana University, Bloomington. Leiden: Brill.
- 2018** Nikolaev, Dmitry and Maria Shkapa. Markirovanije aktantov dvukhmestnykh predikato-  
tov v irlandskom jazyke (munsterskij dialekt) [Participant indexing in bivalent predicates in  
Irish (Munster dialect)] In: ed. Sergey Say. *Valentnostnyje klassy dvukhmestnykh predika-  
tov v raznostruktornykh jazykakh* [Valency classes of bivalent predicates in structurally di-  
verse languages]. St. Petersburg: Institute for Linguistic Studies RAS. 150–167.

## Preprints

- 2021** Choshen, Leshem, Matanel Oren, Dmitry Nikolaev and Omri Abend, *SERRANT: A syntactic classifier for English grammatical error types*. arXiv preprint arXiv:2104.02310

## Published datasets and software

- 2021** Choshen, Leshem, Matanel Oren, Dmitry Nikolaev and Omri Abend. *SERRANT: A syntactic classifier for English Grammatical Error Types* <https://github.com/matanel-oren/serrant>
- 2021** Choshen, Leshem, Dmitry Nikolaev, Yevgeni Berzak and Omri Abend. *SErCl: A cross-lingual taxonomy of syntactic errors* [https://github.com/borgr/GEC\\_UD\\_divergences](https://github.com/borgr/GEC_UD_divergences)
- 2021** Nikolaev, Dmitry. *IPA parser and query engine* <https://github.com/macleginn/eurphon-parse-search>
- 2020** Grossman, Eitan, Elad Eisen, Dmitry Nikolaev and Steven Moran. *SegBo: A database of borrowed sounds in the world's languages*. <https://doi.org/10.5281/zenodo.3832072>
- 2020** Nikolaev, Dmitry, Ofir Arviv, Taelin Karidi, Neta Kenneth, Veronika Mitnik, Lilja Maria Saeboe and Omri Abend. *An aligned subset of Parallel Universal Dependencies*. <https://github.com/macleginn/exploring-clmd-divergences>
- 2020** Goldwasser, Orly, Haleli Harel and Dmitry Nikolaev. *iClassifier: A platform for scientific analysis of classifier scripts and languages*. <https://www.iclassifier.pw/>
- 2018** Nikolaev, Dmitry, Andrey Nikulin and Anton Kukhto. *The Database of Eurasian Phonological Inventories*. <https://eurphon.info/>
- 2018** Nikolaev, Dmitry. *A formant-based vowel classifier*. <https://eurphon.info/static/vowelclf/>

## Invited presentations

- 2021** Koptjevskaja-Tamm, Maria and Dmitry Nikolaev. Warm-hearted people and cold smiles: how universal is AFFECTION AS WARMTH across the languages of the world? *Language, Semantics and Cognition: Saying and conceptualizing the world from Ancient Egypt to Modern Times*. 16–18 April 2021. Yale University
- 2020** Say, Sergey and Dmitry Nikolaev. Launch of BivalTyp. Typological database of bivalent verbs and their encoding frames ([www.bivaltyp.info](http://www.bivaltyp.info)). *Seventeenth Conference on Typology and Grammar for Young Scholars*. 21 November 2020. Institute for Linguistic Studies RAS

## Refereed presentations

- 2020** *Co-occurrence classes of consonants and the modular structure of consonant inventories*. 53rd Annual Meeting of the Societas Linguistica Europaea. 26–29 August 2020. Online
- 2020** *Defining and operationalizing ‘borrowability’ in phonology*. 53rd Annual Meeting of the Societas Linguistica Europaea. 26–29 August 2020. Online (together with Elad Eisen, Steven Moran and Eitan Grossman)
- 2019** *Sound changes and synchronic alternations: A comparative survey*. 13th Conference of the Association for Linguistic Typology. 4–5 September 2019. University of Pavia
- 2019** *The typology of phonological segment borrowing*. 13th Conference of the Association for Linguistic Typology. 4–5 September 2019. University of Pavia (with Eitan Grossman, Elad Eisen, Júda Ronén, Layla Schwartz, Yoav Yosef and Steven Moran)
- 2019** *Eurasian consonant systems through time*. 52nd Annual Meeting of the Societas Linguistica Europaea (a workshop session on Phonological (In)stability and Language Evolution). 21–24 August 2019. Leipzig University
- 2019** *How different were phonological distributions? The World Survey of Phonological Segment Borrowing and the Uniformitarian Assumption*. 52nd Annual Meeting of the Societas Linguistica Europaea. 21–24 August 2019. Leipzig University (with Eitan Grossman and Steven Moran)
- 2018** *Phonological micro-areality in Greater Kurdistan*. 51st Annual Meeting of the Societas Linguistica Europaea. 29 August – 1st September 2018. Tallinn University (with Eitan Grossman)
- 2018** *A distributional typology of valence orientation in Eurasia*. 51st Annual Meeting of the Societas Linguistica Europaea. 29 August – 1st September 2018. Tallinn University (with Eitan Grossman)
- 2018** *Quantifying the extrapolation error of the Family Bias Method*. 4th Usage-Based Linguistics Conference. 2 July 2018. Tel Aviv University (with Victor Kuperman)
- 2017** *Predicting consonant inventories based on areal data*. 12th Conference of the Association for Linguistic Typology. 10–14 December 2017. Australian National University
- 2017** *Weighted lexical semantic maps for areal lexical typology: Verbs of perception and cognition as a case study*. 12th Conference of the Association for Linguistic Typology. 10–14 December 2017. Australian National University (with Thanasis Georgakopoulos, Eitan Grossman and Stéphane Polis)

- 2017 *Towards a distributional typology of sound change in Eurasia: Directionality and areality of sound changes in the IE family*. 50th Annual Meeting of the Societas Linguistica Europaea. 10–13 September 2017. University of Zurich (with Eitan Grossman)
- 2016 *An update on the phonology of Gaelige Chorca Dhuibhne*. 9th Celtic Linguistics Conference. 1–2 September 2016. Cardiff University (with Anton Kukhto)
- 2015 *Tibetan dialects as a case study in areal phonology*. The 48th International Conference on Sino-Tibetan Languages and Linguistics. 21–23 August 2015. University of California, Santa Barbara
- 2011 *Absolute and conjunct endings in Old Irish: A survey of typological parallels*. XIVth International Congress of Celtic Studies. 1–5 August 2011. University of Maynooth, Ireland (with Maria Shkapa)

## Teaching experience

- 2019 *Celtic languages: A typological perspective*, BA course, Hebrew University of Jerusalem
- 2018 *High dimensional and spatial analyses for linguistics*, master-class, University of Liège
- 2017 *A practical introduction to statistics for linguists*, BA course, Hebrew University of Jerusalem
- 2017 *Phonological typology*, MA course, Hebrew University of Jerusalem (co-taught with Eitan Grossman)
- 2012–2016 *Indo-European antiquities*, MA course, Russian State University for the Humanities
- 2014 *Introductory statistics and programming for linguists*, MA course, Moscow State Pedagogical University (co-taught with Maria Shkapa)
- 2013 *Introductory statistics and programming for linguists*, MA course, Moscow State University for the Humanities (co-taught with Maria Shkapa)
- 2012–2013 *Introduction to Old Irish linguistics*, BA course, Lomonosov Moscow State University

## Grants

- 2014 A grant from the Foundation for Fundamental Research in Linguistics for the creation of the Database of Eurasian Phonological Inventories.

## Professional memberships

Association for Linguistic Typology, Societas Linguistica Europaea, Association for Computational Linguistics

## Languages

**Native** Russian

**Fluent** English

**Proficient** French, Hebrew

**Reading** German

**Fieldwork** Irish, Manange

## Technical skills

- Statistical software and deep-learning frameworks (R, stan, pandas, scikit-learn, PyTorch, Julia)
- NLP software packages (NLTK, spaCy)
- General-purpose programming and markup languages (Python, JavaScript, Go, Rust, Java, C/C++, Lisp/Scheme, HTML/CSS,  $\text{\LaTeX}$ ); frontend and backend libraries and frameworks (Flask, Mithril.js, jQuery); basic Unix system administration
- Database-management systems (SQLite)