Understanding Language Evolution Using an Event-Based Model

David M. Goldstein[1,+], **Shawn H. McCreight**[2,+] and **John P. Huelsenbeck**[3,+]

[1] Department of Linguistics
  dgoldstein@humnet.ucla.edu
[2] Nytril LLC
  shawn.mccreight@gmail.com
[3] Department of Integrative Biology
  johnh@berkeley.edu
[+] these authors contributed equally to this work

Bibliography

1. Lewis, P. O. (2001) A likelihood approach to estimating phylogeny from discrete morphological character data . (vol. 50pp. 913..925)

2. Wong, K. M. , Suchard, M. A. & Huelsenbeck, J. P. (2008) Alignment uncertainty and genomic analysis. (vol. 319pp. 473..476)

3. Lunter, G. A. , Miklós, I. & Song, Y. S. (2003) An efficient algorithm for statistical multiple alignment on arbitrary phylogenetic trees . (vol. 10pp. 869..889)

4. Thorne, J. L. , Kishino, H. & Felsenstein, J. (1991) An evolutionary model for maximum likelihood alignment of DNA sequences . (vol. 33pp. 114..124)

5. Chang, W. , Cathcart, C. A. , Hall, D. P. & Garrett, A. J. (2015, March) Ancestry-constrained phylogenetic analysis supports the Indo-European steppe hypothesis . (vol. 91pp. 194..244)

6. Bouchard-Côté, A. , Hall, D. , Griffiths, T. L. & Klein, D. (2013) Automated reconstruction of ancient languages using probabilistic models of sound change . (vol. 110pp. 4224..4229)

7. Holden, C. J. (2002) Bantu language trees reflect the spread of farming across sub-Saharan Africa . (vol. 269pp. 793..799)

8. Bowern, C. & Atkinson, Q. D. (2012, December) Computational phylogenetics and the internal structure of Pama-Nyungan . (vol. 88pp. 817..845)

9. Lartillot, N. & Philippe, H. (2006, April) Computing Bayes factors using thermodynamic integration . (vol. 55pp. 195..207)

10. Wikipedia (2020) Distinctive Feature . Wikipedia, . en.wikipedia.org/wiki/Distinctive_feature

11. Metropolis, N. , Rosenbluth, A. W. , Rosenbluth, M. N. , Teller, A. H. & Teller, E. (1953) Equation of state calculations by fast computing machines . (vol. 21pp. 1087..1092)

12. Jukes, T. H. & Cantor, C. R. (1969) Evolution of protein molecules . *Mammalian Protein Metabolism* (pp. 21..123)Academic Press, .

13. Blevins, J. (2004) Evolutionary phonology . Cambridge University Press , .

14. Wikipedia (2019) Extended Speech Assessment Methods Phonetic Alphabet . Wikipedia, . en.wikipedia.org/wiki/X-SAMPA

15. InternationalPhoneticAlphabet.org (2019) IPA Symbols Chart Complete . InternationalPhoneticAlphabet.org, . www.internationalphoneticalphabet.org/ipa-charts/ipa-symbols-chart-complete

16. Xie, W. , Lewis, P. O. , Fan, Y. , Kuo, L. & Chen, M. (2010, December) Improving marginal likelihood estimation for Bayesian phylogenetic model selection . (vol. 60pp. 150..160)

17. Ringe, D. A. , Warnow, T. & Taylor, A. (2002, March) Indo-European and computational cladistics . (vol. 100pp. 59..129)

18. Wikipedia (2019) International Phonetic Alphabet . Wikipedia, . en.wikipedia.org/wiki/International_Phonetic_Alphabet

19. Straka, G. (1956) La dislocation linguistique de la Romania et la formation des langues romanes à la lumiére de la chronologie relative des changements phonétiques . (vol. 20pp. 249..267)

20. Bloomfield, L. (1933) Language. H. Holt and Company , .

21. Gray, R. D. & Jordan, F. M. (2000) Language trees support the express-train sequence of Austronesian

expansion. (vol. 405pp. 1052..1055)

22. Gray, R. D. & Atkinson, Q. D. (2003) Language-tree divergence times support the Anatolian theory of Indo-European origin. (vol. 426pp. 435..439)

23. Phillips, B. S. , Honeybone, P. & Salmons, J. C. (2015) Lexical diffusion in historical phonology . *The Oxford handbook of historical phonology* (pp. 359..373)Oxford University Press , .

24. Bouckaert, R. R. , Lemey, P. , Dunn, M. , Greenhill, S. J. , Alekseyenko, A. V. , Drummond, A. J. , Gray, R. D. , Suchard, M. A. & Atkinson, Q. D. (2012, August) Mapping the origins and expansion of the Indo-European language family . (vol. 337pp. 957..960)

25. Hastings, W. K. (1970) Monte Carlo sampling methods using Markov chains and their applications . (vol. 57pp. 97..109)

26. Koutna, O. , Niederehe, H. & Koerner, K. E. (1990) On the history of classifications in the Romance language group. *History and historiography of linguistics. Proceedings of the Fourth International Conference on the History of the Language Sciences (ICHoLS IV), Trier, 24..28 August 1987.* (vol. 1pp. 287..300)John Benjamins , .

27. Max Planck Institute for the Science of Human History (2019) PHOIBLE 2.0. Max Planck Institute for the Science of Human History , . www.phoible.org

28. Ohala, J. J. (2003) Phonetics and historical phonology . *The handbook of historical linguistics* (pp. 669..686)Wiley-Blackwell, .

29. Höhna, S. , Landis, . , Heath, . , Boussau, . , Lartillot, . , Moore, . , Huelsenbeck, J. P. & Ronquist, . (2016) RevBayes: Bayesian phylogenetic inference using graphical models and an interactive model-specification language . (pp. 65:726-736)Systematic Biology , . www.revbayes.com

30. Tavaré, S. (1986) Some probabilistic and statistical problems on the analysis of DNA sequences . (vol. 17pp. 57..86)

31. Chen, M. Y. & Wang, W. S. (1975) Sound change. (vol. 51pp. 255..281)

32. Lunter, G. A. , Drummond, A. J. , Miklós, I. & Nielsen, R. (2004) Statistical Alignment: Recent Progress, New Applications, and Challenges . *Statistical methods in Molecular Evolution* Springer Verlag, .

33. Wichmann, . , Søren, . , Holman, E. W. & Brown, C. H. (2018) The ASJP Database (version 18) . ASJP, . asjp.clld.org

34. Malkiel, Y. (1978) The classification of romance languages . (vol. 31pp. 467..500)

35. Weiss, M. (2015) The comparative method. *The Routledge handbook of historical linguistics* (pp. 127..145)Routledge, .

36. Ohala, J. J. (2012) The listener as a source of sound change . *The initiation of sound change* (pp. 21..36) John Benjamins, .

37. Gojobori, T. (1988) To be determined.

Figure 1. For the 'Basic Rules' model, states were grouped into five sets: Nasal Vowel (1), Vowel (2), Nasal Consonant (3), Non Sylabic Sonorant (4) and Consonant (5). Here, the area of the circles is proportional to the estimated equilibrium frequencies for each group. The width of the arrows is proportional to the estimated rates. Note that rates are higher for transitions from one word segment to another when the word segments are in the same group.
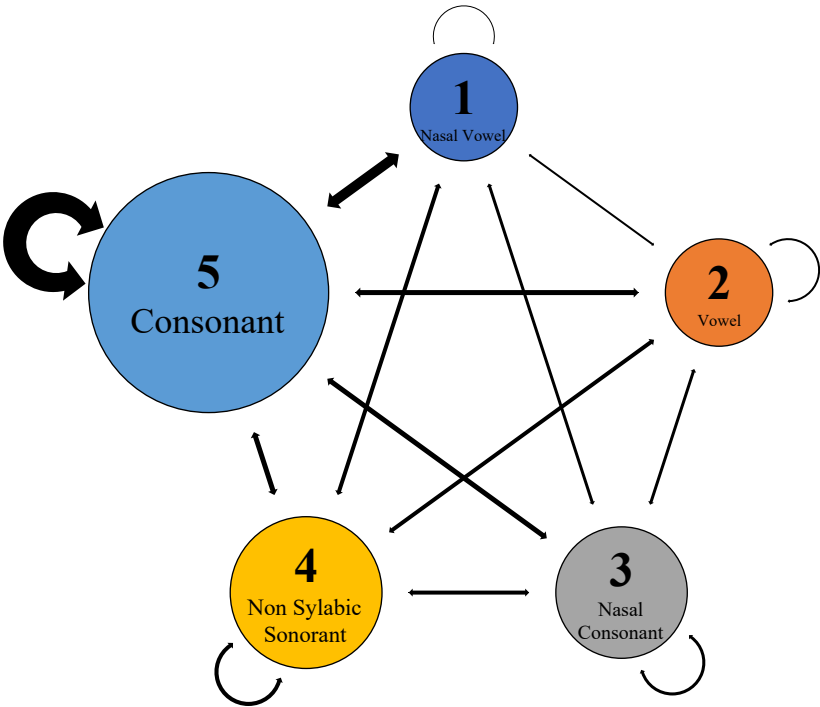
Figure 2. Frequency of occurance of segments in the lexicon
(Max Planck Institute for the Science of Human History , 2019)