

Linguistic convergence and its drivers in Finnish dialects

Jenni Santaharju, Terhi Honkola, Perttu Seppä, Kaj Syrjänen, Unni Leino & Outi Vesakoski

University of Turku, University of Helsinki, University of Helsinki, University of Tampere, University of Tampere, University of Turku

While linguistic divergence increases differences between language varieties, linguistic convergence reduces differences between them (Hinskens et al. 2009; Braunmuller & House 2009). Both processes are essential for understanding the development of present-day language diversity. However, so far linguistic convergence has obtained less attention than linguistic divergence, possibly due to a lack of straightforward measures of linguistic convergence.

Here we adopted a method from population genetics to measure convergence: In biology the method is used to estimate gene flow between populations but here we estimate instead “linguistic flow” between Finnish municipalities. The linguistic data is achieved from Dialect Atlas of Finnish presenting dialectal landscape of Finnish circa 100 years ago (Kettunen 1940).

We first defined dialectal structure of Finnish with model based clustering algorithms (STRUCTURE and BAPS) and assessed how much each hierarchical level – division to two main dialects or 16 regional dialects or dialects within municipalities – explain of the variation in the data. Next, we measured genealogical similarity between the 16 regional dialects by estimating a fixation index (Φ ; see also Syrjänen et al. 2016, Honkola et al. 2018). Finally we estimated linguistic flow

between the dialects by using the method suggested by Tang et al. (2009)

Our main aim was to study the relative importance of geographical distance (measured as km between the centroids of the municipalities) and genealogical similarity (fixation index ϕ) on convergence of the dialects (measured as linguistic flow). For this we run Maximum Likelihood Population Effect models (MLPE).

Our results show that 20 % of the linguistic variation in Finnish 100 yrs ago was contributed to the main dialect division between Eastern and Western dialects. Instead, the linguistic differences between the 16 regional dialects corresponded to 36 % of the variation. However, only 6 % of the variation in linguistic flow between the pairs of 16 regional dialects was due to differences between both-east, both-west and mixed dialect pairs. Instead, the best fitting model included both geographical distance and genealogical similarity and explained 58 % of the variation. Interestingly, the shared genealogical history described a bigger portion of the convergence between the Finnish linguistic variation than sole geographical distance. The framework brings new horizons for empirically testing hypotheses on the drivers of linguistic divergence and convergence.

References

- BRAUNMÜLLER, KURT AND JULIANE HOUSE (2009), Introduction. *Convergence and Divergence in Language Contact Situation*. ed. by Braunmüller, Kurt and Juliane House, 241, Amsterdam: John Benjamins Publishing Company, Online: <https://doi.org/10.1075/hsm.8>
- HINSKENS, FRANS; PETER AUER AND PAUL KERSWILL (2009), The study of dialect convergence and divergence: Conceptual and methodological considerations. *Dialect change: Convergence and divergence in European languages*. ed. by Auer, Peter; Frans Hinskens and Paul

Kerswill. 1-48. Cambridge: Cambridge University Press. Online:
<https://doi.org/10.1017/CBO9780511486623.003>.

HONKOLA, T., RUOKOLAINEN, K., SYRJÄNEN, K., LEINO U, TAMMI, I,
WAHLBERG, N. & VESAKOSKI, O. (2018) Evolution within a language:
Environmental differences contribute to divergence of dialect groups.
BMC Evolutionary Biology. 18:132.

KETTUNEN, LAURI (1940), *Suomen Murteet III A. Murrekartasto*. Helsinki:
Suomalaisen Kirjallisuuden Seura.

SYRJÄNEN, K., HONKOLA, T., LEHTINEN, J., LEINO, A. & VESAKOSKI, O. (2016)
Applying population genetic approaches within languages: Finnish
dialects as linguistic populations”, *Language Dynamics and Change* 6:
235-283.

TANG, JING; WILLIAM P. HANAGE; CHRISTOPHE FRASER AND JUKKA CORANDER. 2009.
Identifying currents in the gene pool for bacterial populations using
an integrative approach. *Plos Computational Biology* 5. Online:
<https://doi.org/10.1371/journal.pcbi.1000455>.