

Evolution of stories in a large-scale online experiment bridging psycholinguistics and cultural evolution



Sébastien Lerique
sebastien.lerique@normalesup.org



Camille Roth
roth@cmb.hu-berlin.de



Abstract

We propose a novel approach for characterizing the short-term evolution of content through transmission chains. These represent a credible model of the iterated transformations that content undergoes when it propagates among individuals. We focus on written utterances, using a Web experiment and a novel analysis to bridge the study of high-level biases with that of psycholinguistics.

Related Work

Currently the study of cultural evolution (Boyd & Richerson 1985, Cavalli-Sforza & Feldman 1981, Sperber 1996) through language is divided between: (1) studies of complex content, restricted to simple or manual analyses, (2) studies of simple content, where complex measures can be used and much is known from psycholinguistics. We propose a descriptive model to link both levels.

Reformulation model

Needleman & Wunsch (1970) extend the Levenshtein distance to align sequences of items

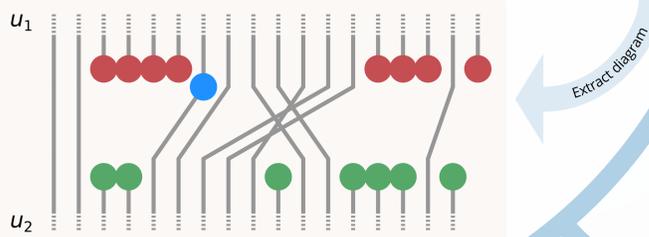
AGAACT → AGAACT-
GACG → -G-AC-G

Extend & apply N.p

At Dover the finale of the - - bailiffs convention Exchange 1 their duty
At Dover - - - - at a Bailiffs convention a speaker said that their duty
said a speaker are delicate dangerous - - - and detailed -
Exchange 1 was to patience and - - - determination

E1 { said a speaker are delicate dangerous E2 }
E2 { a speaker - - - said that }

E2 { said - }
E2 { said that }



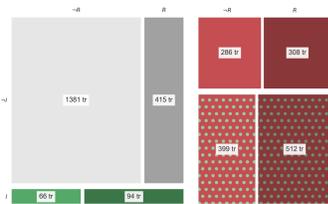
Extract diagram

We leave exchanges aside, obtain sequences of I/R/D, with parent-child correspondences

parent C C D D D D R C C C C C C D D D C D
child C C I I R C C C C I C C I I I C I

Individual operations

Deletions gate other operations

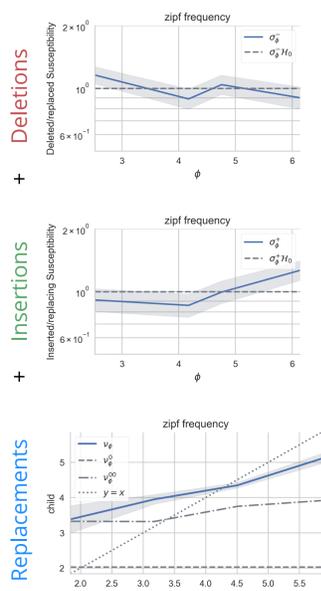


Deletions & Insertions are bursty

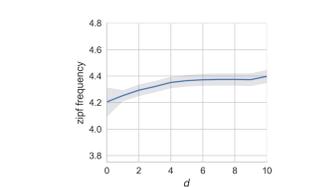


$$B = \frac{\sigma_{intervals} - \mu_{intervals}}{\sigma_{intervals} + \mu_{intervals}}$$

Word features (here word frequency)

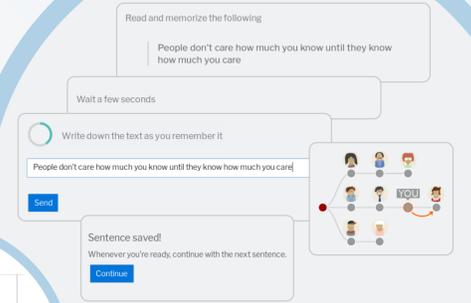


Branch evolution



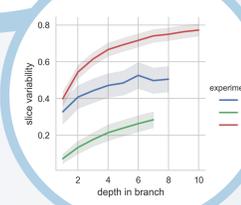
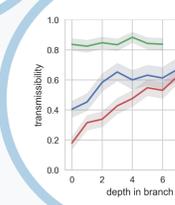
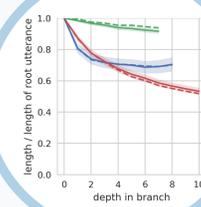
Reformulation trees

gistr.io

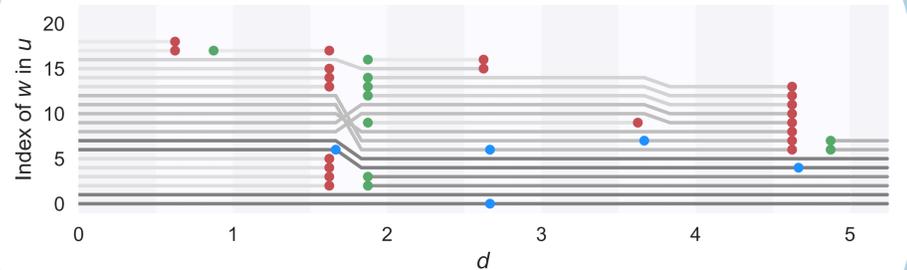


140 participants
50 seeds
1 tree = 7 branches x depth 10

Complex content, complex transformations



Reformulation deep alignments are then stacked to obtain the lineage of each word. The overall branch diagram describes the evolution both in detail and globally.



Line darkness: duration of word survival
Red dots: deletions
Green dots: insertions
Blue dots: replacements

Conclusion

- Using a combination of NLP and biological sequence alignment, we decompose complex transformations in quality data
- Our model identifies key behaviours of transformations at the word and chunk levels
- Results are consistent with in vivo studies
- The model links psycholinguistics with linguistic cultural evolution, a first step forward to integrate levels and improve the parsimony of hypotheses

References

- Boyd & Richerson. The origin and evolution of cultures. New York, N.Y.; Oxford: Oxford University Press, 2005.
- Cavalli-Sforza & Feldman. Cultural transmission and evolution: a quantitative approach. Princeton, NJ, USA: Princeton University Press, 1981.
- Claidière, Scott-Phillips, & Sperber. "How Darwinian is cultural evolution?" In: Philosophical Transactions of the Royal Society B: Biological Sciences 369.1642 (2014), p. 20130368.
- Claidière & Sperber. "The role of attraction in cultural evolution". In: Journal of Cognition and Culture 7.1 (2007), pp. 89-111.
- Kim Sterelny. "Cultural evolution in California and Paris". In: Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences 62 (2017), pp. 42-50.
- Lerique & Roth. "The semantic drift of quotations in blogspace: a case study in short-term cultural evolution". Cognitive Science, 42(1) (2018), pp. 188-219.