Frederik Hartmann
University of Konstanz

RUSE 2019

Investigating the phonological predictability of sound change using deep neural networks

Abstract submission RUSE 2019

Abstract

The traditional view of sound changes being quasi-arbitrary events that do not adhere much to rules and patterns that could be predictable has been frequently challenged and although directionality and long-term conditioning of sound change have been studied and successfully demonstrated [1, 2, 3, 4, 5, 6], not much research has been devoted to analyzing if the phonological conditions of sound changes is trans-temporally applicable, i.e. that the circumstances leading to past sound changes can be extrapolated to correctly predict future changes. In this study, I train deep neural networks on sound changes that occurred from Old High German (OHG) to Middle High German (MHG) and task them afterwards with predicting the sound changes that occurred from MHG to New High German (NHG) which can, in turn, be checked whether they are accurate. The data for this study are 1482 German lemmas that appear in all three language stages considered here which were extracted from the English Wiktionary .xml dump on 20.10.2018 [7]. After processing, the data consisted of 22230 feature matrices encoding each phonetic segment of these lemmas according to 28 phonetic features for each sound in each word and a label vector where binary values indicate whether the particular sound belonging to the respective feature matrix has changed from OHG to MHG. Similarly, the same data matrices and label vectors were created for the change from MHG to NHG. That the features of the sound in question are themselves informative about the occurrence of a change is nothing surprising given that there is a raw statistical likelihood of change by the frequency distribution of changes alone. In the data used in this analysis, OHG /r/ appears in 539 instances and only in 16 of these, it was either reduced or underwent a sound change. This, while being in accordance with previous findings of the changes /r/ did or did not undergo, yields a change probability of P = 0.03 for OHG /r/. These probabilities are detected by a neural network (model I)[1], leading to a prediction accuracy of 0.755 (ROC-AUC score)[2] for changes from MHG to NHG. Yet this finding is unremarkable since it only shows a deep neural network can detect conditional probabilities for sound-specific change events and that these probabilities are temporally consistent enough to map on later changes in the language's history fairly well. To investigate whether sound changes are predictable via a word's particular phonological circumstances, it is necessary to train a network to detect whether a change occurs from OHG to MHG given *only* the phonological environment of the sound in OHG. If this network is, in turn, able to predict changes from MHG to NHG correctly on the basis of the OHG-MHG data, it would be evidence for trans-temporal phonological conditioning of sound changes. This can be done with two different approaches: (1) I trained a model on the full OHG sound environment data to detect change-prone environments independent of the respective sound's features (model II). The results show a prediction accuracy of 0.760 (ROC-AUC score) for changes from MHG to NHG. The problem with this approach is that it is not sound-specific and thus only considering *general* features of change-prone environments. Moreover, we encounter the risk of the network, in reality, inferring the target sound features and then estimating the probabilities similar to the network mentioned above. (2) A solution to this problem would be to train separate networks on single sounds where we can be sure the change probability is the same for all samples during training since it is the same for all samples. In doing this, we could force the network to actually 'learn' conditions for sound changes. Yet to analyze every sound individually, the sample sizes for individual sounds is too small with too few changes to safely train a neural network on. In preliminary tests on some of these few samples, the networks performed between 5 and 10 percent above the random baseline, however to obtain reliable results, further research is needed. The overall results show that sound changes are at least somewhat predictable from the conditions of sound changes present in the previous language stage. This observation supports the claim that sound changes are not singular events but can be predicted with some accuracy from phonological conditions only.

[1]For further details on all trained models, please refer to the appendix.
[2]With 50 percent expected by a random baseline.

Appendix

Table 1: Network architecture and evaluation for model I

Network type	Dense neural network	
Output type	Binary	
Optimizer	Adam	
Batch size	250	
Layer	Layer size	Activation
Dense layer 1	64	ReLU
Dense layer 2	32	ReLU
Output layer	2	softmax

Pred. results test data	Model evaluation	Random baseline
Precision	0.388	0.110
Recall	0.685	0.491
F1 score	0.495	0.181
ROC-AUC score	0.773	0.481
Pred. results MHG-NHG		
Precision	0.388	0.1
Recall	0.658	0.483
F1 score	0.447	0.165
ROC-AUC score	0.755	0.491

Table 2: Network architecture and evaluation for model II

Network type	Multi-input CNN	
Output type	Binary, weighted	
Optimizer	Adam	
Batch size	250	
Layer	Layer size	Activation
Conv. layer 1a and 1b	256	ReLU
Conv. layer 2a and 2b	128	ReLU
Dense layer 1a and 1b	64	ReLU
Dense layer 2	256	ReLU
Dense layer 3	128	ReLU
Output layer	2	softmax

Pred. results test data	Model evaluation	Random baseline
Precision	0.316	0.114
Recall	0.882	0.496
F1 score	0.466	0.185
ROC-AUC score	0.818	0.498
Pred. results MHG-NHG		
Precision	0.259	0.101
Recall	0.775	0.489
F1 score	0.388	0.168
ROC-AUC score	0.760	0.496

References

[1] Patricia J. Donegan and Geoffrey S. Nathan. Natural phonology and sound change. In Patrick Honeybone and Joseph Curtis Salmons, editors, *The Oxford handbook of historical phonology*, Oxford handbooks in linguistics, pages 431–449. Oxford University Press, Oxford, 2015.

[2] Juliette Blevins. Evolutionary phonology: A holistic approach to sound change typology. In Patrick Honeybone and Joseph Curtis Salmons, editors, *The Oxford handbook of historical phonology*, Oxford handbooks in linguistics, pages 485–500. Oxford University Press, Oxford, 2015.

[3] J. J. Ohala. Coarticulation and phonology. *Language and speech*, 36 (Pt 2-3):155–170, 1993.

[4] J. J. Ohala. The phonetics of sound change. In Charles Jones, editor, *Historical linguistics*, Longman linguistics library, pages 237–278. Longman, London, 1993.

[5] Mark Hale. Neogrammarian sound change. In Brian D. Joseph, editor, *The handbook of historical linguistics*, Blackwell handbooks in linguistics, pages 343–368. Blackwell, Malden, MA, 2003.

[6] Ricardo Bermúdez-Otero. Individual differences and the explanation of sound change, 2017.

[7] Wiktionary, the free dictionary. https://en.wiktionary.org. Accessed: 2019-03-13.