Outline

Exemplar Approach to Vowel Harmony

VH as an evolutionary process

- vowel harmony is the result of a diachronic process of sound change motivated by perturbatory effects of coarticulation on listener perception [Ohala, 1994, Blevins, 2004]
- evidence suggests it is possible to extract phonological rules corresponding to VH from coarticulatory patterns [Przezdziecki, 2005]
    - counter-claim: phonetic basis for VH irrelevant from a synchronic point of view [Anderson, 1980, Nevins, 2010]
- motivates search for models that integrate phonetic features and phonological structures

Motivating the exemplar approach

- mental lexicon integrates detailed representations of acoustic/auditory, articulatory, visual, spatial, and even social inputs
- phonological structures are implicitly learnt through repeated exposure to *exemplars* of these inputs
- exemplar models have been argued to be particularly suited to modelling language change as an evolutionary process
- Wedel [2006], Port [2007], Johnson [2007], Coleman [2002], Johnson [2006]

Connectionist models (a.k.a. Neural Networks)

Connectionist models in linguistics

- linked to psycholinguistically motivated accounts of production/perception, and processing [Dell et al., 1999, Hawkins and Smith, 2001, Port, 1990]
- proposed to be compatible with emergent/exemplar accounts of phonological acquisition [Bybee and McClelland, 2005, Lathroum, 1989, Hare, 1990, Rodd, 1997, Cole, 2009, Alderete and Tupper, 2018, Cole, 2009]
- but, actual early implementations were often limited; toy models, sketches

Where does this experiment come in?

- the success of *“deep learning”* [Manning, 2015]
- it is now possible to train such models to learn semantic representations directly from raw, unannotated phonetic or text inputs
- resurgence of interest in testing if and how the representations learnt by such *“end-to-end”* models maps to traditional linguistic structures [Alishahi et al., 2017, Doucette, 2017, Gulordava et al., 2018, Ravfogel et al., 2018, van Schijndel and Linzen, 2018, Enguehard et al., 2017, Linzen et al., 2016]

Where does this experiment come in?

proposed aim:

- replicate methodology from Alishahi et al. [2017] to test the representation of VH structures in an RNN model trained in an exemplar-compatible way

why recurrent?

- recurrence allows processing sequential inputs (like speech)
- lets the network learn interactions between different portions of the input
    - sensitive to coarticulatory effects

not in scope:

- testing the biological plausibility of RNNs (*short answer:*

Research Questions, Data, and Methods

Research Questions

RQ1

- Can structures corresponding to vowel harmony classes be found through the vector representations learned by a *weakly-supervised* “deep” RNN

RQ2

- If yes (to RQ1), then what is the hierarchical position of the layer(s) that best learns to represent these structures w.r.t. the rest of the network
  - Do the results of Alishahi et al. [2017] replicate?

Dataset

- Aalto University DSP Course Conversation Corpus
- Finnish spontaneous conversations with force-aligned transcriptions (utterance, word, segment)
- 5200 utterances, 9.7hrs of audio from 218 male and 24 female speakers
- Finnish has backness harmony, proceeding left-to-right
- `http://urn.fi/urn:nbn:fi:lb-2017092133`

Model

Input

- Mel-frequency cepstral coefficient (MFCC) vectors, utterance level

Layers

- 1 convolutional, 5 recurrent

Output

- Embedding layer, trains to project utterance encoding and image encoding to a joint vector space

Experiment

ABX discrimination task

- tuples of the form (A,B,X) where
  - A, B, and X are CV syllables
  - B and X vowels share harmonic class
- for each tuple, calculate $sign(dist(A, X) - dist(B, X))$, where $dist(i, j)$ is euclidean distance between model's vector representations of syllables i and j
- +ve sign indicates that the model has learned to discriminate the VH classes in A and B

Feedback

Questions? Comments? Suggestions?

- project repository at
  `https://git.irfus.in/irfan/ExemplarRNNHarmony.git`

References

References

J. Alderete and P. Tupper. Connectionist approaches to generative phonology. *The Routledge Handbook of Phonological Theory. Routledge*, 2018.

A. Alishahi, M. Barking, and G. Chrupała. Encoding of phonology in a recurrent neural model of grounded speech. In *Proceedings of the 21st Conference on Computational Natural Language Learning (CoNLL 2017)*, page nil, - 2017. doi: 10.18653/v1/k17-1037. URL `https://doi.org/10.18653/v1/k17-1037`.

S. R. Anderson. Problems and perspectives in the description of vowel harmony. *Issues in Vowel Harmony*, page 1, 1980. ISSN 0165-7763. doi: 10.1075/slcs.6.02and. URL http://dx.doi.org/10.1075/slcs.6.02and.

J. Blevins. Evolutionary phonology. 2004. doi: 10.1017/cbo9780511486357. URL http://dx.doi.org/10.1017/cbo9780511486357.

J. Bybee and J. L. McClelland. Alternatives to the combinatorial paradigm of linguistic theory based on domain general principles of human cognition. *The Linguistic Review*, 22(2-4): nil, 2005. doi: 10.1515/tlir.2005.22.2-4.381. URL https://doi.org/10.1515/tlir.2005.22.2-4.381.

J. Cole. Emergent feature structures: harmony systems in exemplar models of phonology. *Language Sciences*, 31(2-3): 144–160, 3 2009. ISSN 0388-0001. doi: 10.1016/j.langsci.2008.12.004.

J. Coleman. Phonetic representations in the mental lexicon. *Phonetics, phonology and cognition*, pages 96–130, 2002.

G. S. Dell, F. Chang, and Z. M. Griffin. Connectionist models of language production: Lexical access and grammatical encoding. *Cognitive Science*, 23(4):517–542, 1999. doi: 10.1207/s15516709cog2304_6. URL https://doi.org/10.1207/s15516709cog2304_6.

A. Doucette. Inherent biases of recurrent neural networks for phonological assimilation and dissimilation. In *Proceedings of the 7th Workshop on Cognitive Modeling and Computational Linguistics (CMCL 2017)*, pages 35–40, 2017.

E. Enguehard, Y. Goldberg, and T. Linzen. Exploring the syntactic abilities of rnns with multi-task learning. *arXiv preprint arXiv:1706.03542*, 2017.

K. Gulordava, P. Bojanowski, E. Grave, T. Linzen, and M. Baroni. Colorless green recurrent networks dream hierarchically. *CoRR*, 2018. URL `http://arxiv.org/abs/1803.11138v1`.

M. Hare. The role of trigger-target similarity in the vowel harmony process. *Annual Meeting of the Berkeley Linguistics Society*, 16(1):140, 1990. doi: 10.3765/bls.v16i0.1724. URL `https://doi.org/10.3765/bls.v16i0.1724`.

S. Hawkins and R. Smith. Polysp: A polysystemic, phonetically-rich approach to speech understanding. *Italian Journal of Linguistics*, 13:99–188, 2001.

K. Johnson. Resonance in an exemplar-based lexicon: the emergence of social identity and phonology. *Journal of Phonetics*, 34(4):485–499, 2006. doi: 10.1016/j.wocn.2005.08.004. URL `https://doi.org/10.1016/j.wocn.2005.08.004`.

K. Johnson. Decisions and mechanisms in exemplar-based phonology. *Experimental approaches to phonology. In honor of John Ohala*, pages 25–40, 2007.

A. Lathroum. Feature encoding by neural nets. *Phonology*, 6 (02):305–316, 1989. doi: 10.1017/s0952675700001044. URL `https://doi.org/10.1017/s0952675700001044`.

T. Linzen, E. Dupoux, and Y. Goldberg. Assessing the ability of lstms to learn syntax-sensitive dependencies. *arXiv preprint arXiv:1611.01368*, 2016.

C. D. Manning. Computational linguistics and deep learning. *Computational Linguistics*, 41(4):701–707, 2015. doi: 10.1162/coli_a_00239. URL `https://doi.org/10.1162/coli_a_00239`.

A. Nevins. *Locality in Vowel Harmony*. EBSCO ebook academic collection. MIT Press, 2010. ISBN 9780262140973. URL `https://books.google.com/books?id=XIG8ZDzuAHcC`.

J. J. Ohala. Towards a universal, phonetically-based, theory of vowel harmony. In *Third International Conference on Spoken Language Processing*, 1994.

R. Port. How are words stored in memory? beyond phones and phonemes. *New Ideas in Psychology*, 25(2):143–170, 2007. doi: 10.1016/j.newideapsych.2007.02.001. URL `https://doi.org/10.1016/j.newideapsych.2007.02.001`.

R. F. Port. Representation and recognition of temporal patterns. *Connection Science*, 2(1-2):151–176, 1990. doi: 10.1080/09540099008915667. URL https://doi.org/10.1080/09540099008915667.

M. Przezdziecki. *Vowel harmony and coarticulation in three dialects of Yoruba: phonetics determining phonology*. PhD thesis, Cornell University, 2005.

S. Ravfogel, F. M. Tyers, and Y. Goldberg. Can lstm learn to capture agreement? the case of basque. *CoRR*, 2018. URL http://arxiv.org/abs/1809.04022v1.

J. Rodd. Recurrent neural-network learning of phonological regularities in turkish. *CoNLL97: Computational Natural Language Learning*, 1997.

M. van Schijndel and T. Linzen. Modeling garden path effects without explicit hierarchical syntax. In *Proceedings of the 40th Annual Conference of the Cognitive Science Society*, pages 2600–5, 2018.

A. B. Wedel. Exemplar models, evolution and language change. *The Linguistic Review*, 23(3):nil, 2006. doi: 10.1515/tlr.2006.010. URL `https://doi.org/10.1515/tlr.2006.010`.

Appendix

VH in Finnish and Bangla

- Finnish: backness harmony, proceeds left-to-right
    - front or back vowel in the initial syllable spreads that feature to vowels in non-initial syllables
    - three harmonic classes - front [ä ö y]; back [a o u]; neutral [e i][1]
    - e.g., pos+ahta+(t)a → posahtaa (back)
    - räj+ahta+(t)a → räjähtää (front)
- Bangla: ATR harmony, proceeds right-to-left
    - ATR feature of suffixal vowels spreads to stem vowels
    - e.g., pɔtr+ika → potrika
    - khɛl+i → kheli

[1] neutral vowels are unchanged by harmony