Background

- Age of exposure is a predictor of second language proficiency
- Native-like phonological proficiency is attained only by learners exposed at the earliest ages (simultaneous bilinguals)
- Does early commitment to phonological patterns account for age of exposure effects?
- Two computational models were trained using unsupervised learning (Toscano & McMurray, 2008; McMurray et al., 2009)
  - 2D Gaussian Mixture Model
  - Hebbian Normalized Recurrence Network

Language data

- Spanish has two front vowels and two back vowels
- Quichua has one front vowel and one back vowel

Guion (2003) studied four groups of Quichua-Spanish bilinguals exposed to Spanish at different ages
- Simultaneous bilinguals distinguished Spanish and Quichua vowels
- Late learners (age of first exposure > 13 yrs) produced only one front vowel and one back vowel
- Learners exposed before age 13 varied in the number of vowels they produced
- Front vowels were more often distinguished than back vowels
- Both models were trained on tokens from distributions modeled on the productions of monolingual speakers

References


Models

2D Gaussian Mixture Model (GMM)
- Categories defined by 2D (F1 x F2) Gaussian distributions
- Each Gaussian defined by 5 parameters:
  - $\phi$: likelihood of category
  - $\mu_1$ and $\mu_2$: mean along each cue dimension
  - $\sigma_1$ and $\sigma_2$: standard deviation along each dimension
  - $\rho$: correlation between F1 and F2
- Each model is initialized with 200 Gaussians with equal $\phi$, $\sigma$, and $\rho$ and random $\mu$ values
- During training, model is presented with F1 and F2 values
- Parameters updated via maximum likelihood estimation and winner-take-all competition (McMurray et al., 2009)
- Trained on 80,000 training trials per run (30 runs total)

Hebbian Normalized Recurrence Network (HNRN)
- Two-layer neural network
- 200 topographically organized input units with Gaussian tuning curves for each dimension (F1 and F2)
- 20 output units corresponding to possible vowel categories
- On each training trial:
  - Pair of F1 and F2 values are presented at input layer
  - Activation feeds forward to output units
  - Output units compete
  - Activation feeds back to input units
  - Weights are updated via Hebbian learning
  - Processes repeats until output unit activation settles
- Trained on 40,000 training trials per run (30 runs total)

Simulation groups

- Monolingual Quichua and Spanish groups
- Simultaneous bilingual: trained on tokens from both languages
- Early bilingual: started training with Quichua and switched to Spanish after 25% of total training time
- Mid bilingual: switched after 50% of total training
- Late bilingual: switched after 75% of total training

Results

- Total number of categories
- Mean F1 and F2 values of model categories
- Number of front vs. back vowel discriminations

Analysis

- Number of categories
- Mean F1 and F2 values of model categories
- Number of front vs. back vowel discriminations

Human speakers


Conclusions

- Both models learn the vowel systems of Quichua and Spanish when trained on monolingual data
- When exposed to bilingual data, both models distinguish different numbers of categories depending on whether L2 (Spanish) is introduced, paralleling data from bilingual speakers with different ages of acquisition
- Age of exposure effect in the models arises as a consequence of learning without any changes in model plasticity (learning rates)
- Language users may be restricted in learning a second language not because of a critical period, but by the morphemes that the system has already made to the first language; a separate mechanism is not necessary

Future directions

- GMM categories are sometimes unstable and with too much training data, the model may end with fewer categories than in the language (unlike human learners); further exploring the parameter space of the model may lead to a more stable set of parameters
- HNRN can be given additional types of input, such as context information or labels for different lexical items; could allow the network to maintain distinctions between morophemes
- Look at whether it is more or less difficult to learn an L2 with more categories than L1 (Quichua->Spanish) or to learn one with fewer categories (Spanish->Quichua)

Acknowledgements

We would like to thank Bob McMurray for helpful discussions about these simulations. This work was supported by a postdoctoral fellowship awarded by the CRLM to M.C. and NIH D008859 awarded to Bob McMurray.

meghan.clayards@mcgill.ca, joseph-toscano@uiowa.edu