Modeling the Collapse of Variation in Pidgin Development

David Ruskin, PhD
University of Rochester

Presented:
SPCL Winter Meeting 2018

UNIVERSITY of ROCHESTER

Historical Trend

- Early pidgin languages tend to be *highly* variable
  (Mühlhäusler, 1979; DeGraff, 1999; Holm, 2000)
- If they creolize …notable reduction in variation
  (Mühlhäusler, 1986; DeGraff, 1999; Adone, 2012)

- Reduction of variability occurs fastest in the **lexicon**; **grammatical** features often remain variable into the creole stage
  (Mühlhäusler, 1979; Sankoff, 1980; Smith, 2002)

Big Question

- Why do different parts of a language lose variability / stabilize at different rates?
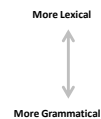

One hypothesis:

- *Maybe individuals' ability to learn and reshape their input drives the historical trend*

- Prior laboratory studies suggest some basic psychological mechanisms may influence stabilization
  (Hudson Kam & Newport, 2005, 2009; Hudson Kam & Chang, 2009)

- But…

Prior Work

- These studies have not looked at stabilization across different grammatical categories
- Therefore, cannot be used to compare against our observed historical trend

- Ruskin (2014)
  - Miniature language study looking at variation in different linguistic categories
  - Would participants' productions mirror the historical trends?

Prior Work: Language

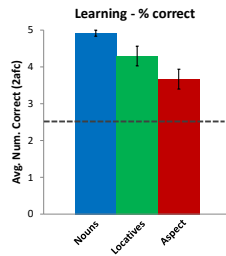- Variability in lang. occurred at three grammatical depths:

  More Lexical
  - Nouns
  - Locational Adjectives / Locative markers
  - Aspect markers
  More Grammatical

- Each noun, etc., had two forms in free variation, differing only by rate of occurrence (Major form: 67%, minor form: 33%)

Two Main Results

1) Participants learned different categories at different rates
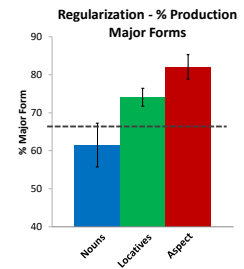   (worse at more grammatical categories)

- Similar to real-life late-learning adults (see Johnson & Newport, 1989)

Learning - % correct

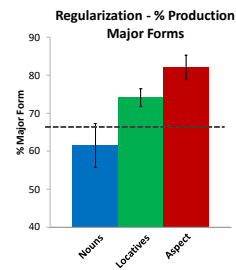(Avg. Num. Correct (zafc): Nouns, Locatives, Aspect)

Two Main Results

2) Participants' own productions showed reduced variability* with increasing grammaticality

*reduction of variability = stabilization or regularization

Regularization - % Production Major Forms

(% Major Form: Nouns, Locatives, Aspect)

But!

- This →
  is backwards from the historical trend!

- Why?...

Regularization - % Production Major Forms

(% Major Form: Nouns, Locatives, Aspect)

Missing factor?

- Experiment – looked at individuals in isolation
- But communication is a *communal* act
- The two basic results will produce different patterns at the population-level
  – Categories, like nouns, that allow variation and are easily learned should spread through the community quickly
  – Other categories should spread more slowly

- But how does variation drop out of the population?

Alignment

- When we talk to someone,
  we are likely to use the same words and structures
- This is a process called **alignment / coordination** / etc.
  (Clark, 1996; Pickering & Garrod, 2004)

- Could radically reshape variation at the community-level
  – Will make aligned variants appear much more frequent
  – While language is forming, this could lead to permanent changes
    (c.f. Trudgill, 2008)

- Maybe this is what was missing!

So….

- Where do we get a community of new speakers?

- It's difficult / ethically impossible to try to run a full-scale experiment
  – Without starting a land war, can't force population shifts

Model

- Can *simulate* a community using agent-based modeling
  (Sun, 2006)
  - Custom built in JavaScript, visuals with d3.js (https://d3js.org/)
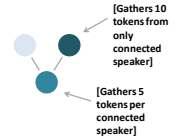
Model basics

← Each circle is a person

Each person communicates with (& learns from) other people they are connected to

Model Basics

- We will track 3 independent, variable features:
  - Nouns (Blue)
  - Loc. Adj. (Green)
  - Aspect (Red)

- Each has two possible forms, e.g.:(A & B) or (present / absent)
- Each individual can use the two forms in any proportion
  - 0% form A (Light Circles)
  - 100% form A (Dark Circles)   ← ~50/50 form A & form B

Learning

- In a given epoch (i.e., on a given turn)
- Each person:
  - Listens
  - Speaks
- Listening: gathers 10 tokens per person
- Speaking: tokens are produced by weighted die rolls

[Gathers 10 tokens from only connected speaker]

[Gathers 5 tokens per connected speaker]

Learning

- All people in the simulation listen and speak at the same time
- And then all update simultaneously

- Updating is by a simple Hebbian learning algorithm

$$P_{t+1} = \lambda * P_{observed} + (1-\lambda) * P_t$$

This individual's probability of using a feature next turn

Learning rate (constant)

Probability of feature from current sample

Individual's current probability of using feature

Setting up the Community

- As each person added,
- Connect them to up to 3 random people already in the community

- Initial proportion of A/B (or +/-) is 67% (like lab experiments)

These are the bare bones

- Haven't added in:
  - Learning rates
  - Regularization
  - Alignment

- First need to see what default behavior is like
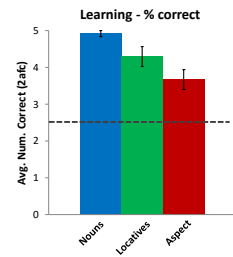
[open in browser]

Adding in factors

- Let's add in:
  - Variable Learning Rates
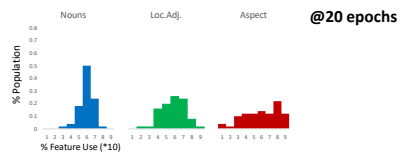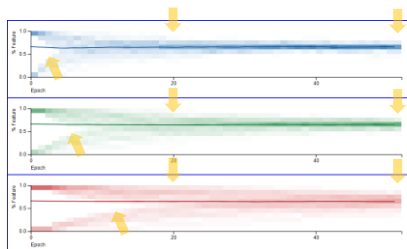 - Regularization
  - Alignment

Learning Rates

- Participants had trouble learning more grammatical features

- Implemented by rates estimated from results:
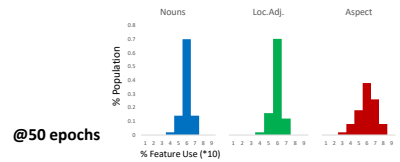 - Locative rate = 2/3 Noun rate
  - Aspect rate = 4/9 Noun rate

Learning - % correct

Avg. Num. Correct (2 aft)

Nouns   Locatives   Aspect

Learning Rate

- 1 Example run, 50 epochs

- Produces different rates of convergence,
- Variation is retained

Nouns   Loc.Adj.   Aspect

@20 epochs

% Population

% Feature Use (*10)

Histograms from single epochs help show the trend

Nouns   Loc.Adj.   Aspect

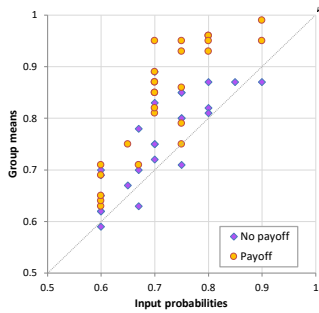@50 epochs
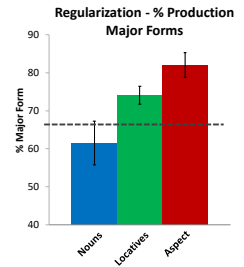
% Population

% Feature Use (*10)

Learning rate alone

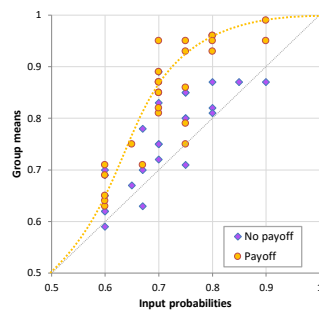- Not enough to get historical trends

- Let's add in regularization

Regularization

- We have regularization rates from the experiments
- But this only gives us rates when the starting ratio was 67% / 33%

Regularization - % Production Major Forms
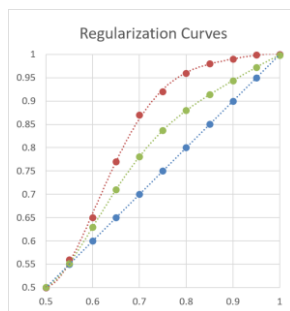

Diagonal:
Input = output

- For the full function, can turn to non-linguistic probability-matching tasks
- Data from: Vulkan (2000)
- Each point is a different paper: average participant regularization

- Regularization rates from lab language studies...
- closely mirror boosting behavior on tasks that reward correct responses
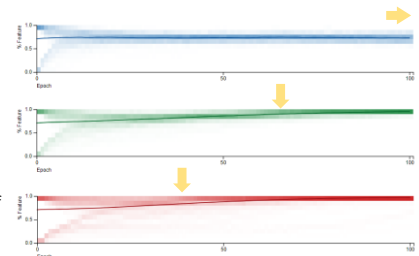- Estimate of function from best fit polynomial

Regularization rates

- To approximate lab data, Loc.Adj. rate is half that of Aspect
- Nouns are not regularized

Regularization Curves

Learning *& Regularization*

- Average of repeated runs
- Aspect stabilizes first, then Loc., Nouns remain variable
- Arrows: Point where majority of population uses only one variant
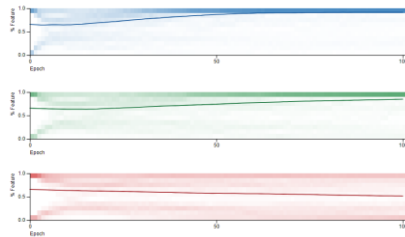
So Far – Reverse of Historical Trend

- This means that merely having a community of speakers is not enough
- Something else is needed

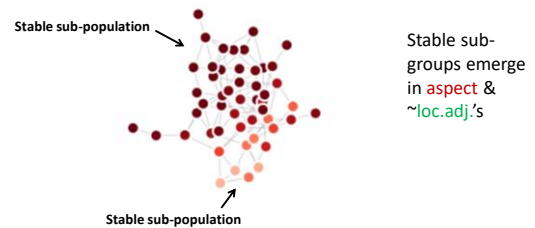- Will adding alignment flip these trends?

Alignment

- Implementation:
- Alignment relies in part on common ground
  (Clark & Wilkes-Gibbs, 1986)

- In the model, if two people *both* favor a given variant, we increase their likelihood of using it that epoch

Learning, Regularization, *& Alignment*

- Average of repeated runs

- Trend has flipped!

- Strong differences in means (solid lines)

Learning, Regularization, *& Alignment*

Stable sub-population

Stable sub-population

Stable sub-groups emerge in aspect & ~loc.adj.'s

Historical trend

- Much of the variation described in creoles is dialectal variation
- The formation of stable sub-groups in the simulations closely mirrors that

Results!

- We simulated a community of speakers that communicate in a very basic way
- And then by adding only a few basic factors:
  - Ease of learning different categories
  - Regularization
  - Alignment
- We were able to reproduce observed historical trends

Bigger picture

- Development of pidgin and creole languages is immensely complex
- But underlying some changes may be very simple psychological and social processes working together

Future Directions

- How do these effects change if the network is sparse / densely connected?
- What's the interaction between regularizing rate and learning rate?
- Other learning algorithms?
- What happens when people are born into the community? Or when we have a lot of late learning adults enter?

Thanks to:

- Advising on dissertation work:
  – Elissa Newport @ Georgetown

- Feedback:
  – Steve Piantadosi, computation & language lab @ Rochester
  – Keturah Bixby @ Rochester

- And you!

Questions?

Contact info

David Ruskin, PhD
University of Rochester
david.ruskin@rochester.edu

slides & demos:
drdrphd.com/creolemodel

UNIVERSITY of ROCHESTER

References

Adone, D. (2012). *The Acquisition of Creole Languages: How Children Surpass Their Input*: Cambridge University Press.

Bock, R.D. (1975) *Multivariate Statistical Methods in Behavioral Research*. New York: McGraw-Hill.

Clark, H. H. (1996). *Using language*. Cambridge university press.

Clark, H. H., & Wilkes-Gibbs, D. (1986). Referring as a collaborative process. *Cognition*, *22*(1), 1-39.

DeGraff, M. (1999). *Language creation and language change: creolization, diachrony, and development*. Cambridge, Mass.: MIT Press.

Hebb, D. O. (1949). *The organization of behavior: A neuropsychological approach*. John Wiley & Sons.

References

Holm, J. (2000). *An Introduction to Pidgins and Creoles*: Cambridge University Press.

Hudson Kam, C. L., & Chang, A. (2009). Investigating the cause of language regularization in adults: Memory constraints or learning effects? *Journal of experimental psychology. Learning, memory, and cognition, 35*(3), 815.

Hudson Kam, C. L., & Newport, E. L. (2005). Regularizing unpredictable variation: The roles of adult and child learners in language formation and change. *Language Learning and Development, 1*(2), 151-195.

Hudson Kam, C. L., & Newport, E. L. (2009). Getting it right by getting it wrong: When learners change languages. *Cognitive Psychology, 59*(1), 30.

References

Johnson, J. S., & Newport, E. L. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology, 21*(1), 60-99.

Mühlhäusler, P. (1979). Growth and Structure of the Lexicon of New Guinea Pidgin (Pacific Linguistics: C-52). *Canberra: The Australian National University*.

Mühlhäusler, P. (1986). *Pidgin & Creole linguistics*. Oxford, OX, UK ; New York, NY, USA: B. Blackwell.

Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and brain sciences*, *27*(2), 169-190.

References

Ruskin, D. (2014). *Cognitive influences on the evolution of new languages.* (PhD), University of Rochester.

Sankoff, G. (1980). *The social life of language*: University of Pennsylvania Press.

Smith, G. P. (2002). *Growing up with Tok Pisin: contact, creolization, and change in Papua New Guinea's national language*: Battlebridge.

Sun, R. (Ed.). (2006). *Cognition and multi-agent interaction: From cognitive modeling to social simulation*. Cambridge University Press.

References

Trudgill, P. (2008). Colonial dialect contact in the history of European languages: On the irrelevance of identity to new-dialect formation. *Language in Society*, *37*(2), 241-254.

Vulkan, N. (2000). An economist's perspective on probability matching. *Journal of economic surveys*, *14*(1), 101-118.