Variation and Social Networks during Language Change

MCWOP Presentation
Oct 21, 2006

Celina Troutman, Brady Clark, Matt Goldrick
Northwestern University

Social Networks and Language Change

• A specific issue that social networks have been used to address:
  › We don’t always talk to everyone near us geographically - we have a network of relationships
    We speak like our close friends
    We pick up new language features through our friends’ other connections
  › Many studies on language change spreading based on social networks (Milroy & Milroy 1985, Milroy 1992)

The Problem

How does an initially rare variant spread through a speech community (Sapir 1921)?

› THRESHOLD PROBLEM (Nettle 1999)

Learners should adopt the language forms they hear the most, so how does this sort of change happen all the time?

Example of overcoming the Threshold Problem

• Denasalization of word medial [ŋ] to [ɡ] in Japanese (Hibiya 1996)
  › About 80 years ago, medial [ɡ] was quite rare
  › Period of variation between [-ŋ-] and [-ɡ-]
  › Today [ɡ] is the dominant form - [ŋ] is hardly used
Example of overcoming the Threshold Problem

Source: Hibiya (1996: 163)

Overview

- Previous studies of the threshold problem have incorporated social networks (Nettle 1999)
- Limitation
  - represented speakers as having categorical grammars - no intraspeaker variation
- Present study
  - model a social network of speakers with probabilistic grammars
- Results - did not solve the threshold problem
- Conclusion - embedding speakers in a social network structure is not sufficient to allow a rare variant to spread

Social Network with Categorical Grammars

- Replicated findings from Nettle (1999) showing that a social network of learners with categorical grammars can solve the threshold problem
- Three major components of the model
  - Language users - who are they?
  - Social network structure - how do they interact?
  - Learning algorithm - how do they change?

Language Users

- Speakers can have one of two types of grammars
  - /ŋ/ > [ŋ]  (Faith(nasal) >> * ŋ)
  - /ŋ/ > [g]  (* ŋ >> Faith(nasal))
- Speakers produce utterances in accord with their grammar
  - People always utter either [ŋ] or [g], but not both
Social Network Structure

As learners enter the network, they want to attach to those people who already have many connections (Albert & Barabási, 1999)

Popularity tends to lead to even greater popularity - a few ‘hubs’

---

Why this network structure?

- Real social networks tend to have a few well-connected items and many less-connected ones (Barabási 2003)
  - Ex. personal relationships, the Internet, academic paper citations
- It is one type of a larger class of networks with shared properties (Barabási 2003)
  - Suggests results found here may be generalized to other network structures

---

Learning Algorithm

At each iteration, language users interact with each other

- **SPEAKING** - speak to a neighbor depending on my grammar
  - If grammar is: /n/ > [ŋ]
  - Then utter [ŋ]
- **LISTENING** - change your grammar to what you heard
  - If you hear [ŋ]
  - Then change your grammar to /ŋ/ > [ŋ]

---

Learning algorithm for Categorical Model

Let’s see the model in action ...
Results from Categorical Model

• Ran networks of 20, 40, or 60 nodes, 12 times each, for 12 iterations
• Convergence within just a few iterations
• Usually converged to the majority variant, but minority could also spread

Could solve the threshold problem!

Limitations of Categorical Model

• Population Size
• Oversimplified network structure
• Oversimplified learning algorithm
• Only two linguistic variants
• Lack of random noise in production or mistakes in perception

No intraspeaker variation

Language Users in the Probabilistic Model

• Speakers can have both types of grammars
  - /n/ > [ŋ] (Faith(nasal) >> * n)
  - /n/ > [g] (* n >> Faith(nasal))
• Each grammar is associated with a weight - the probability of accessing that grammar
• Speakers produce utterances in accord with the grammar accessed
  - Individuals have a probability of uttering either [ŋ] or [g]

Intraspeaker variation during language change

Source: Hibiya (1996: 163)
Learning Algorithm in the Probabilistic Model

Linear reward / penalty algorithm:
(Bush & Mosteller 1951, 1958; Yang 2002)

- SPEAKING - choose a grammar based on their probability weights. Then speak to a neighbor using that grammar:
- If chosen grammar is: /ŋ/ > [ŋ]
- Then utter [ŋ]
- LISTENING - update the probability of using grammar based on what was heard:
  - If you hear [ŋ]
  - Then increase your probability of using grammar /ŋ/ > [ŋ]

More on the Probabilistic Model

- Distinction between proportion of use of a variant at population level and individual level
  - Percentage of all people using [ŋ] at one iteration
  - Probability of one person using [ŋ] at one iteration
- Only looking at the special situation where a small percentage of population initially uses one grammar (G1) categorically, the rest uses G2 categorically
  - Alternatively, a small percentage to could initially use G1 variably, and the rest use G2 categorically

Results from Probabilistic Model

- Ran networks of 20, 40, or 60 nodes, 12 times each, for 8000 iterations
- Initially, 25% of the population has grammar outputting [ŋ-]
- Everyone agrees on a grammar relatively quickly (within ~300 iterations)

Threshold problem is unsolved!

Results from Probabilistic Model

- Language users do not converge on 0 or 1
- They agree on some intermediate grammar state and continue to vary together
Summary

- **Problem** - how does an initially rare variant spread through a population?
- **Potential Solution** - embed speakers in social networks
  - Limitation - does not account for intraspeaker variation
- **Potential Solution** - simulate speakers with probabilistic grammars
- **Result** - does not solve the threshold problem
- **Conclusion** - embedding speakers in a social network structure is not sufficient to allow a rare variant to spread

Future Directions

- **Conditions for solving the threshold problem with categorical grammars**
  - Effect of initial percentage of population with the rare variant, specific network structure, etc.
- **Restructuring network over time**
  - Aging, "critical period" for learning
- **Speaker feedback**
  - Learn from your own utterances as well as from others
  - May have an effect on 'entrenchment' - gradual decrease in an individual's category variability over time (Pierrehumbert 2000)

Acknowledgements

- Janet Pierrehumbert and Uri Wilensky for feedback on an earlier version of this project
- Robert Daland for many helpful discussions
- The Ling 500 class for their questions & comments

References


