City University
of Hong Kong

Language Death:
Modeling Bilingualism & Social Structure

James W. Minett
Language Engineering Laboratory
City University of Hong Kong

Workshop on Exploratory Models of Language Acquisition
Santa Fe Institute
October 20, 2003

Language Death

- A language dies when an entire generation of speakers fails to transmit the language to the next generation (if any).
- Krauss (1992) estimates that **90%** of the world's 6,000 or so languages will die within **100** years.
- More "conservative" estimates put the figure at **50%** (Crystal, 2000).
- Pagel (1995) estimates ***c*. 140,000 languages** have ever existed (median); so the majority of the world's languages are already dead.

Language Death

- Distribution of numbers of first language speakers in the World:

NUMBER OF SPEAKERS:	*LANGUAGES*	*CUMULATIVE %*
≥ 100,000,000	8	
10,000,000 – 99,999,999	72	99.9%
1,000,000 – 9,999,999	239	98.7%
100,000 – 999,999	795	94.8%
10,000 – 99,999	**1,605**	**81.6%**
1,000 – 9,999	**1,782**	**55.1%**
100 – 999	1,075	25.7%
10 – 99	302	8.0%
1 – 9	181	3.0%
	6,059	

Table adapted from (Crystal, 2000); original data from *Ethnologue* (February 1999).

Language Death

- So although the **language faculty** as a whole is extremely robust, **individual languages** are not.

- **What do we want to know?**

  In a situation in which two or more languages compete for speakers:

  – how **likely** is an endangered language to die?

  – how can this likelihood be **reduced**?

- **Aim:**

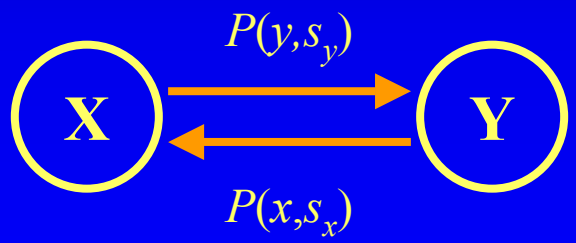
  To model the **dynamics of language death** and strategies for **maintaining** endangered languages

Dynamical Systems Model — Abrams & Strogatz (2003)

- **Assumptions**:
  - Two languages, X & Y, **compete** for speakers
  - All speakers **monolingual**
  - **No** language **change**
  - Social structure and age distribution assumed **homogeneous**

  - Speakers **acquire** a language based on its **attractiveness**
  - Attractiveness is a monotonic function of **status** & **number of speakers**

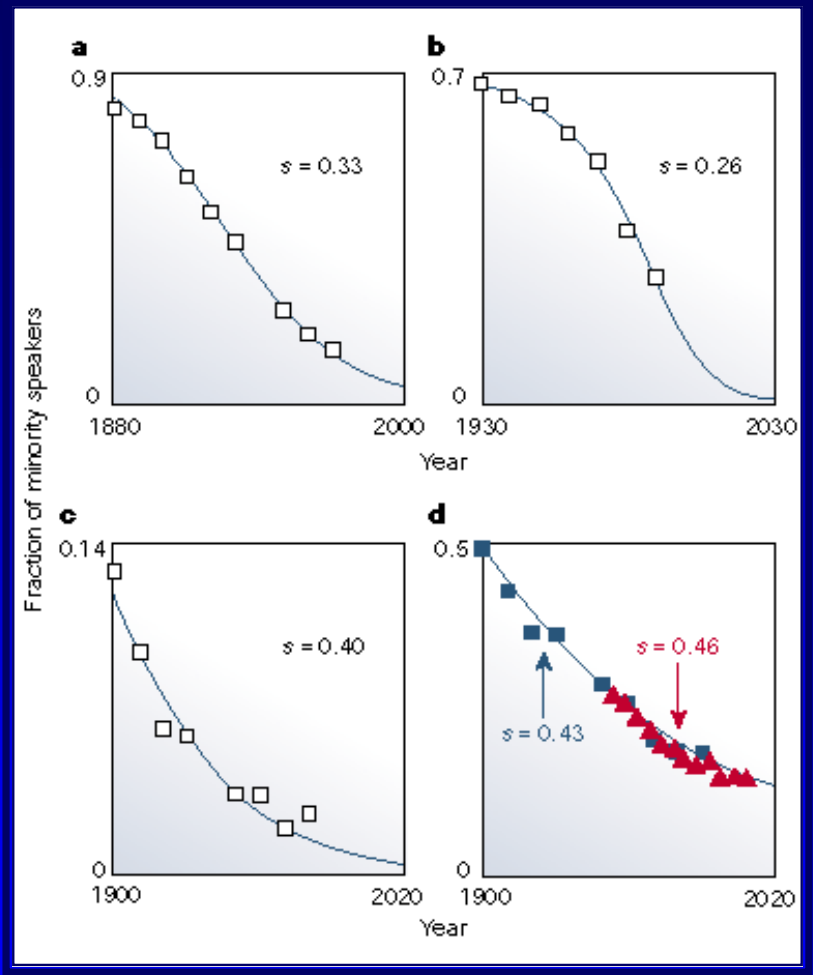
- **Dynamics**:



$$P(x, s_x) = c\, x^a\, s_x$$

Dynamical Systems Model — Abrams & Strogatz (2003)

- **Endangered languages studied:**
  - a) **Scottish Gaelic** (Sutherland, Scotland)
  - b) **Quechua** (Huanuco, Peru)
  - c) **Welsh** (Monmouthsire, Wales)
  - d) **Welsh** (all Wales)
- **Prediction:**
  one language will **acquire all** speakers;
  the other language will **die**
- Despite its **simplicity**, this model fits the data very well
- But the model does not account for **bilingualism**



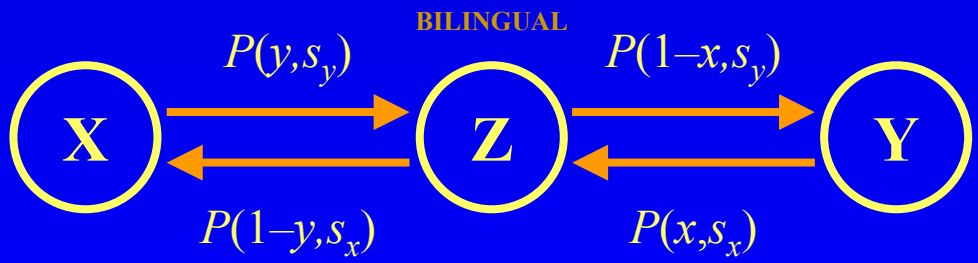
Abrams, D. M. & Strogatz, S. H. 2003.
Modelling the dynamics of language death. *Nature* **424**:900.

Modeling Bilingualism

- **Assumptions**:
  - Two languages, X & Y, **compete** for speakers
  - **Speakers may be monolingual or bilingual (Z)**
  - **No** language **change**
  - Social structure and age distribution assumed **homogeneous … for now**

  - Speakers **acquire** a language based on its **attractiveness**
  - Attractiveness is a **monotonic** function of **status** & **number of speakers**

- **Dynamics**:

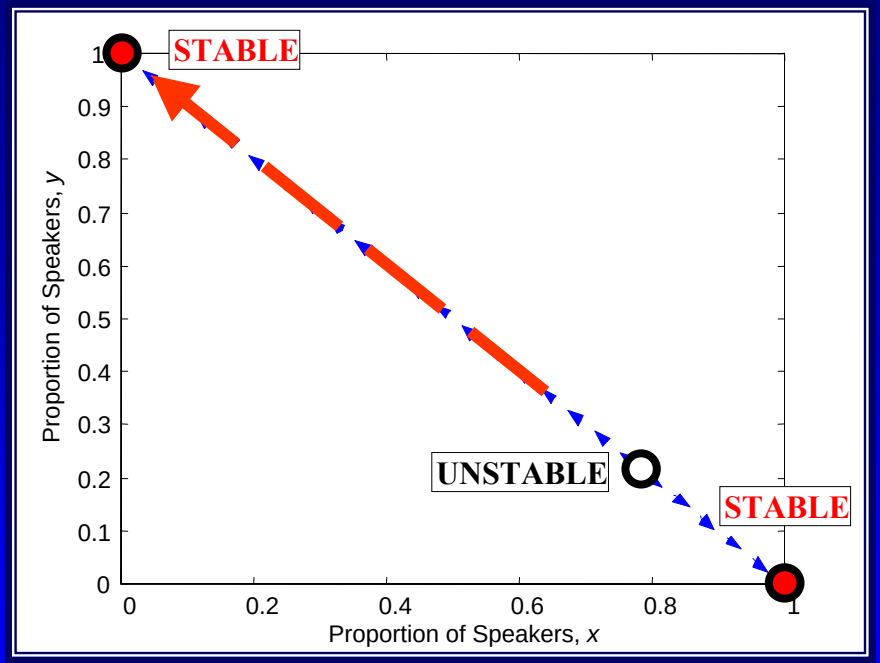


$$P(x, s_x) = c\, x^a\, s_x$$

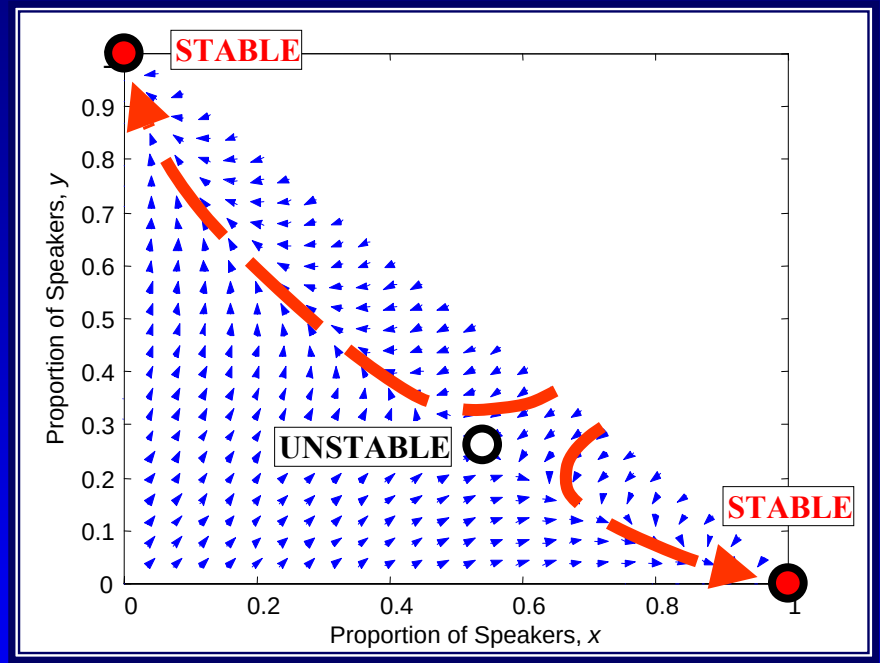
Dynamics of the Models

- **Phase portraits:**

Abrams & Strogatz Model



Our Model



- According to both models, **language death** is **inevitable** …

… but **is it**?

Language Maintenance

- According to Crystal (2000), an **endangered** language can be maintained by:
  - increasing the **prestige** of its speakers;
  - increasing the **wealth** of its speakers;
  - increasing the **power** of its speakers;

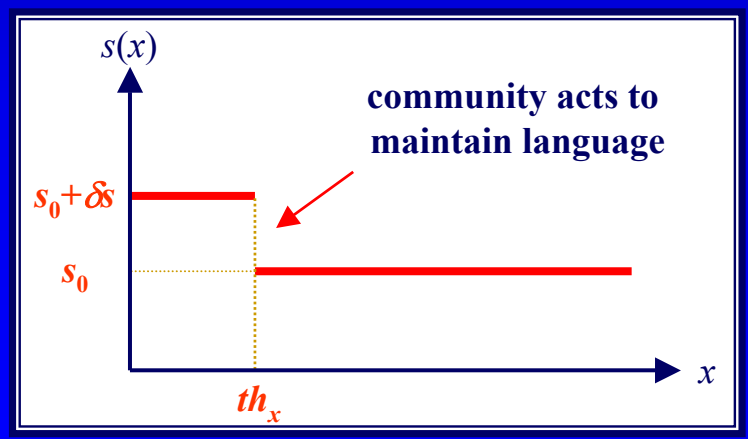
  } **status**

  - improving presence in **educational system**;
  - insuring that the language can be **written down**;
  - providing access to **electronic technology**

  } **access**

- Abrams & Strogatz suggest maintaining a language by **control** on its **status**

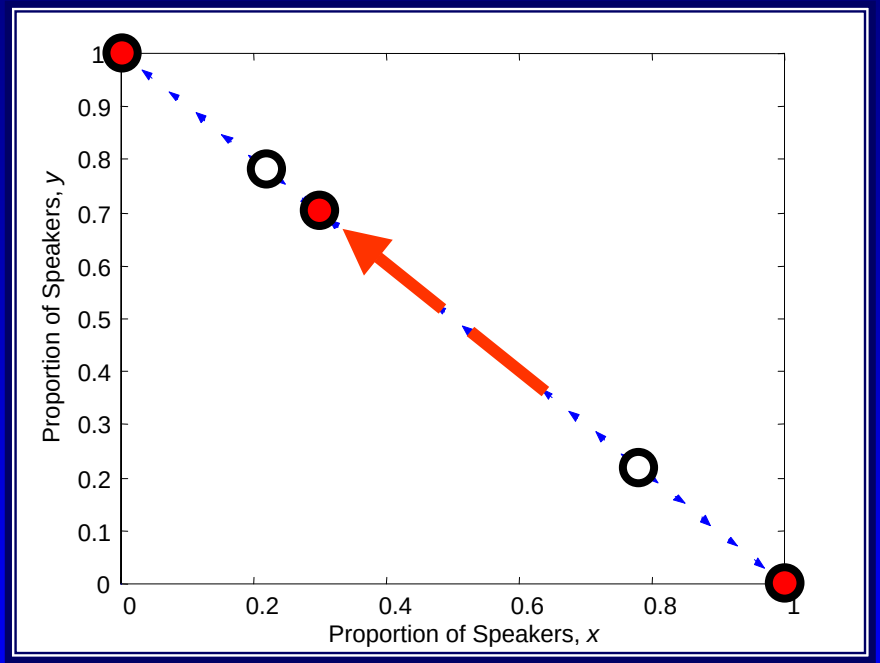
- Here we model **relative status** (i.e. prestige/wealth/power) with the function:



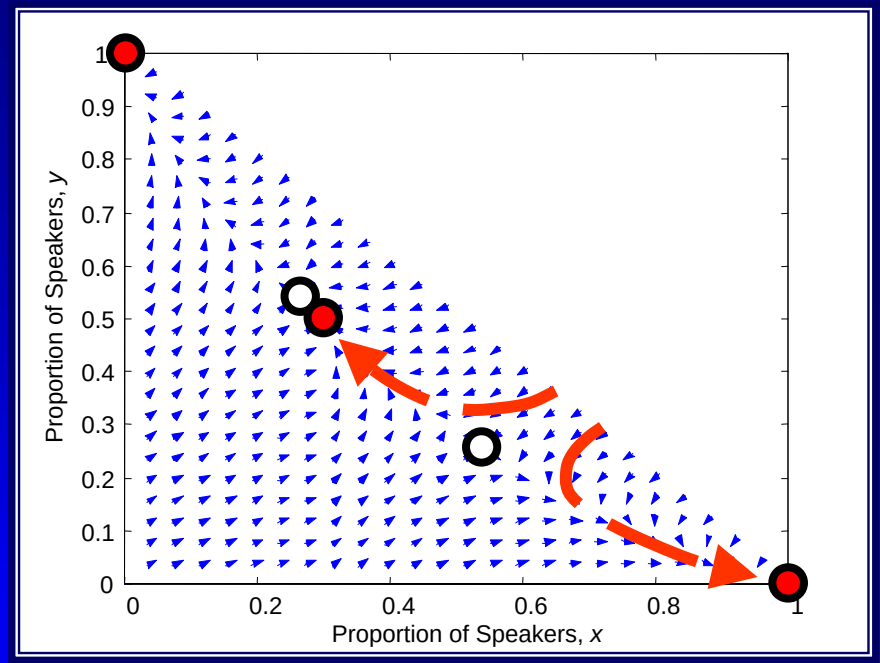
Dynamics of the Models with Language Maintenance

- Phase portraits for **language maintenance** (a=1.3, s=0.4, ds=0.2, th_x=0.3):

Abrams & Strogatz



Our Model

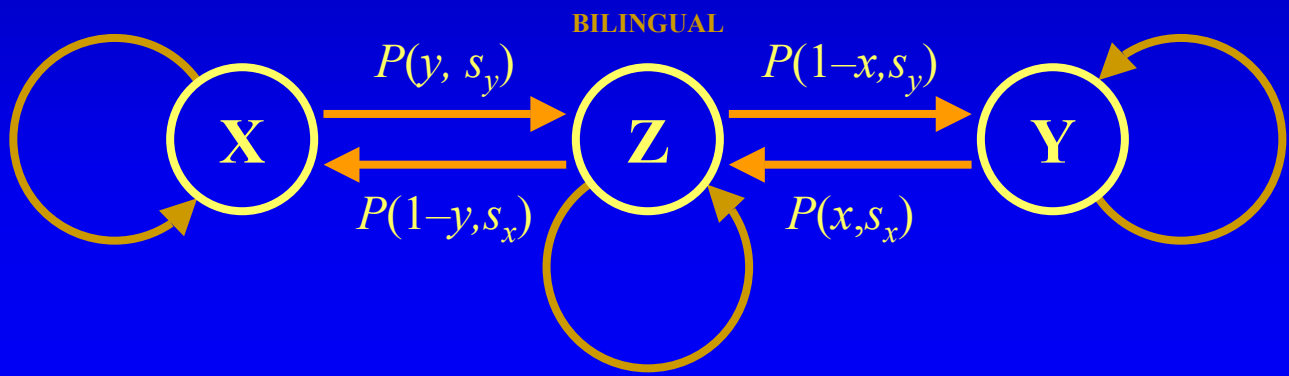


- According to both models, **both languages** can be **maintained** for **some** initial conditions …

… but **how often**?

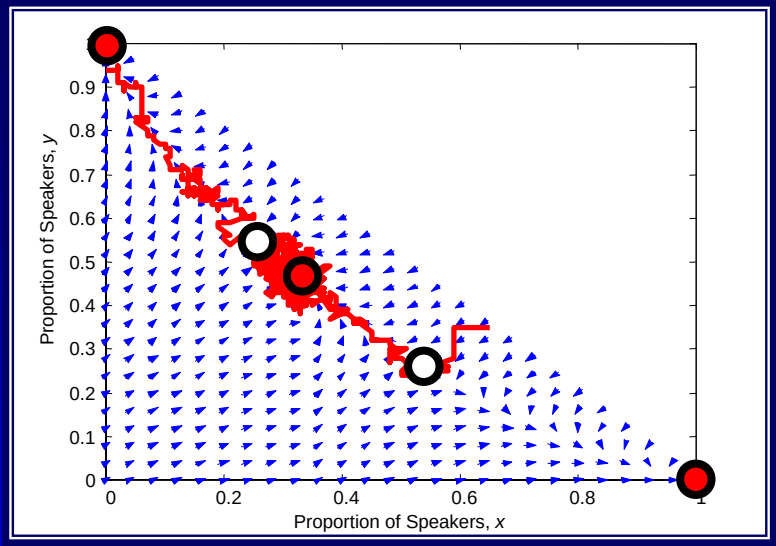
Stochastic Nature of Language Death

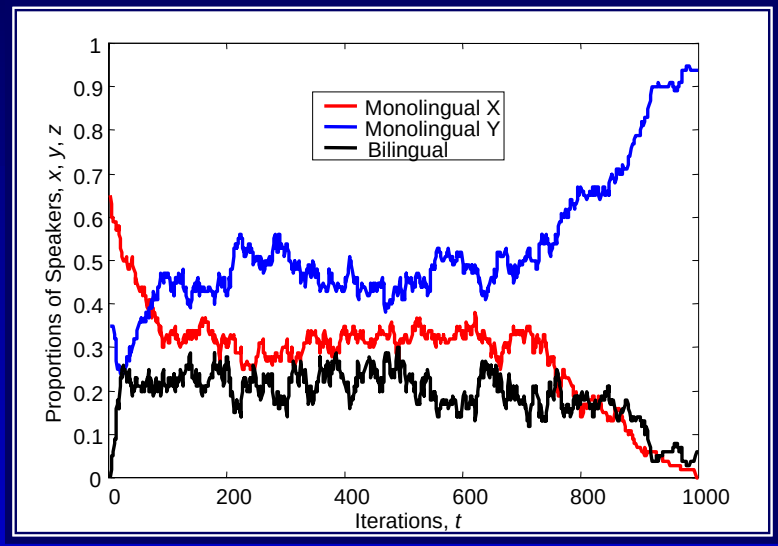
- The above models are deterministic
- But we cannot predict the language(s) acquired by a particular speaker
- So we model the system as a multi-agent system / Markov process
- Agents change states iteratively according to the transition probabilities:

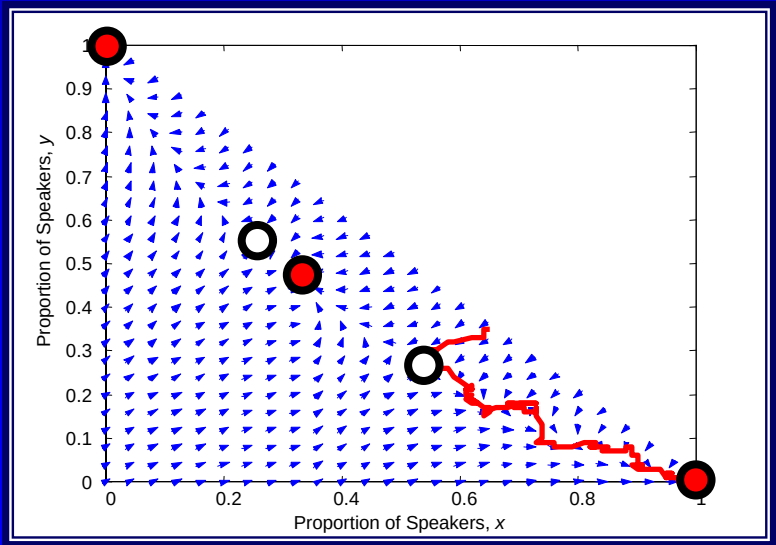


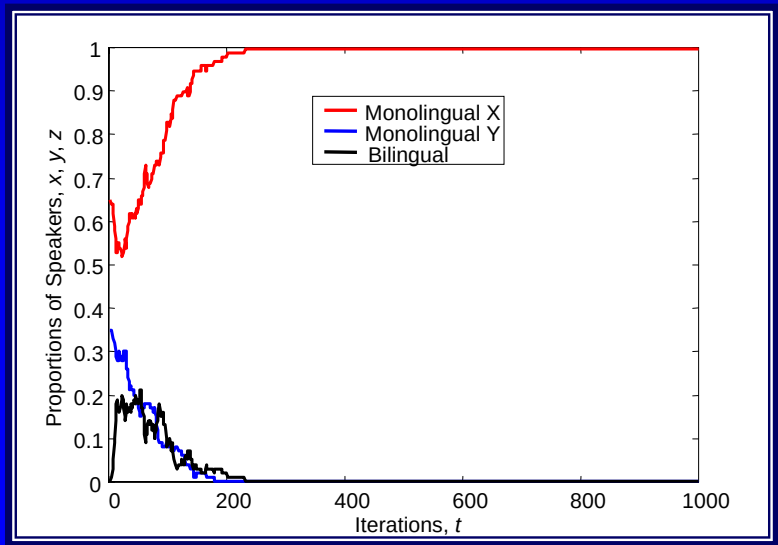
Stochastic Behavior — Our Model

(x_0=65%, a=1.3, s=0.4, δs=0.2, th_x=⅓)



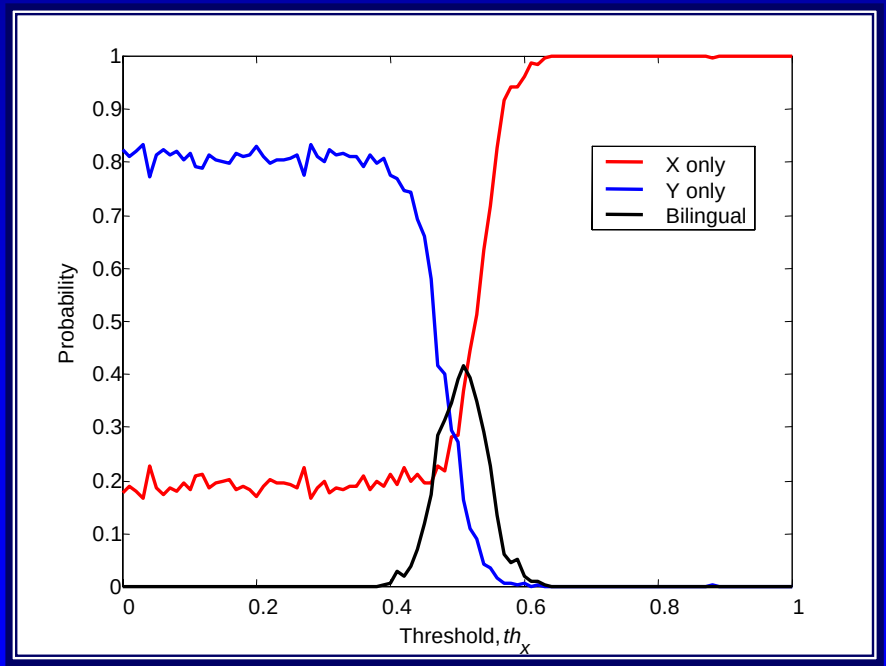




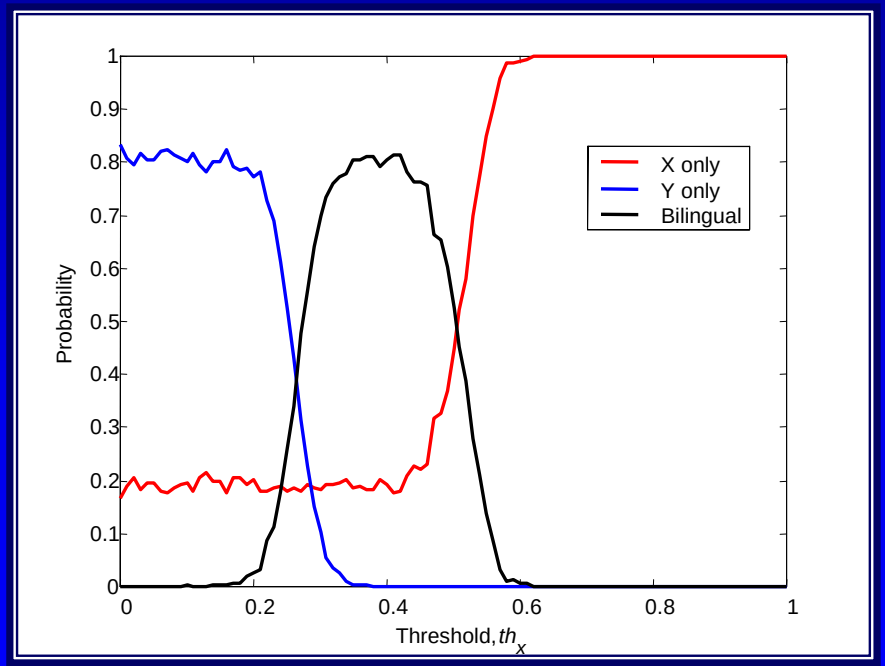


So What Does This Tell Us ?

- We can predict the **probability** that an endangered language can be **saved** after action is taken to maintain it:
  - **by how much** should the **status** of the endangered language be **increased**?
  - **how soon** should the community **act** to maintain an endangered language?



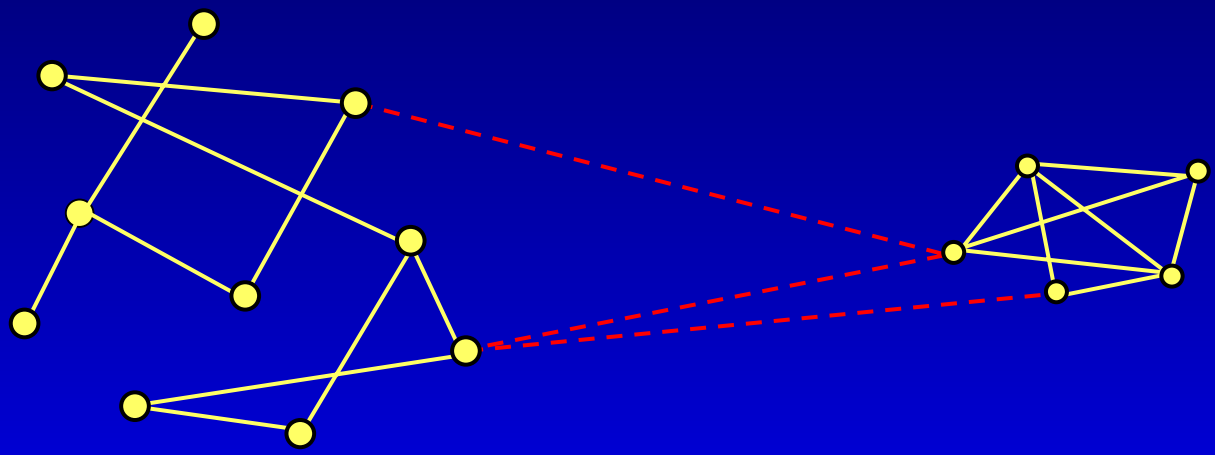
(x_0=75%, a=1.0, s=0.4, δs=**0.1**)



(x_0=75%, a=1.0, s=0.4, δs=**0.2**)

But speakers do not interact exhaustively!

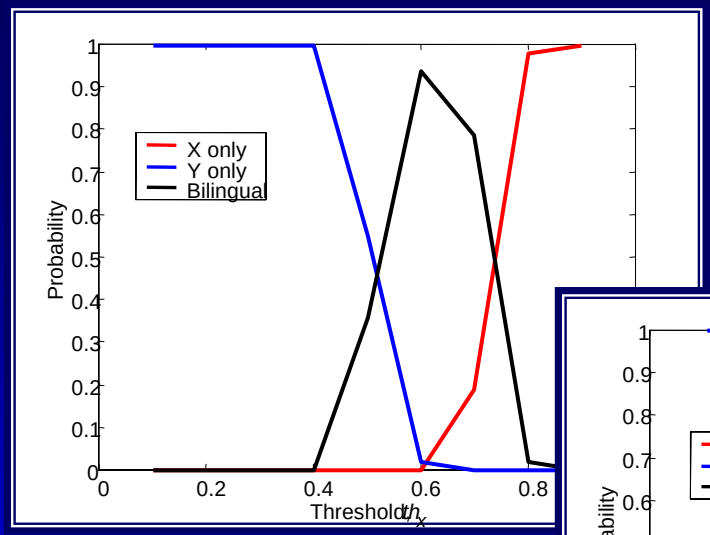
- Model **heterogeneous social structure** using **local-world networks** (Li & Chen, in press), for example:



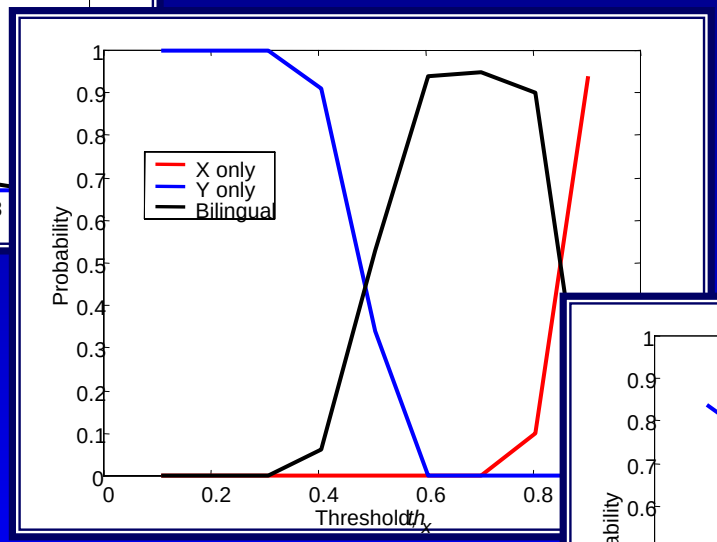
- What role does the structure of the **social network** play in the maintenance of an endangered language?

- Is niche construction relevant here?

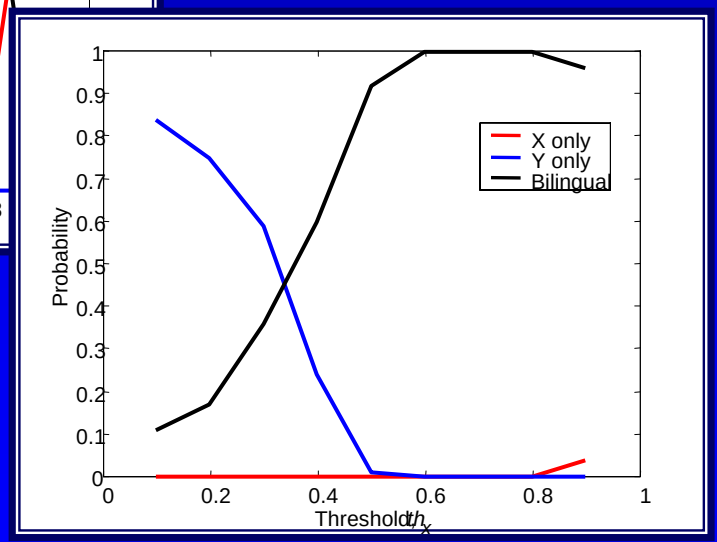
Effect on Probability of Maintenance

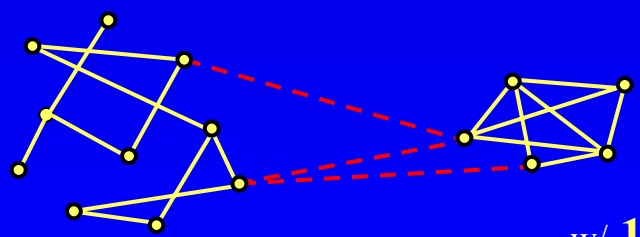


(x_0=75%, a=1.0, s=0.3, δs=0.2)



w/ **5%** inter-connectivity





w/ **1%** inter-connectivity

Preliminary Findings

- In fully connected social networks, endangered languages can be maintained by **enhancing the status** of the language within some window:
  - status must be **sufficiently enhanced**;
  - status must be enhanced **before** the endangered language becomes **moribund**, but **not so soon** that the **competing language** itself becomes **endangered**.

- The language maintenance strategy proposed here is **not robust**.

- All other factors being equal, endangered languages spoken in **isolated communities** appear to be **more easily maintained**.

Future Work

- Improve modeling of **social structure**:
  - **preferential linkage** to agents with a **common language**
  - **evolution of social structure**
  - model **age distribution**
  - build networks to better **fit actual social networks**
- **Collect data**, both monolingual & **bilingual**, for endangered languages:
  - plan to work with minority languages of China, **Bai** & **She**
- **Fit models** to the **data**
- **Compare** the fit of this **bilingual** model to the **monolingual** model of A&S
- Relationship with other **diffusion** systems, including lexical diffusion

Bibliography

- **Crystal, D.** (2000) *Language Death* (CUP: Cambridge).
- **Nettle, D. & Romaine, S.** (2000) *Vanishing Voices: the extinction of the world's languages* (OUP: New York).
- **Ethnologue**, http://www.ethnologue.com. **Krauss, M.** (1992) "The world's languages in crisis." *Language* 68, 4–10.
- **Pagel, M.** (1995) Contribution to the Conservation of Endangered Languages seminar, University of Bristol, 21 April 1995. *Iatiku* 1, 6.
- **Abrams, D. M. & Strogatz, S. H.** (2003) "Modelling the dynamics of language death." *Nature* 424, 900.
- **Watts, D. J.** (1999) *Small Worlds: The Dynamics of Networks Between Order and Randomness* (Princeton University Press: Princeton).
- **Barabasi, A. L. & Albert, R.** (1999) "Emergence of scaling in random networks." *Science* 285:5439, 509–512.
- **Li, X. & Chen, G.** (in press) "A local-world evolving network model." *Physica A* (in press).