Testing transmission and diffusion with an agent-based model

Laurence Kenny and James N. Stanford (*Dartmouth College*)

When Baxter et al. (2009) published an LVC article modeling the emergence of New Zealand English, they exemplified the growing importance of computational/mathematical modeling in variationist sociolinguistics. Following their work, we use "agent-based modeling" (ABM) to test Labov's (2007) notion of transmission and diffusion. In ABM, simple interactions between individual "agents" can lead to the emergence of complex patterns (Blythe & Croft 2009; Harrison et al. 2002). In this study, (1) we first tested the validity of our model by comparing it to real-world dialectology, and then (2) modeled the chain-shift effects of Labov (2007). Hypothesis: Simple density of interactions among agents, combined with a "blank slate" model of child dialect acquisition, is sufficient to account for the logarithmic curves found in dialectometry (Nerbonne 2010) and the chain-shift observations in Labov (2007). (1) Real-world dialectometry in Germany, Norway, U.S., etc., finds a logarithmic relationship between geographic distance and dialect differences (Kretzschmar et al. 2010; Heeringa & Nerbonne 2001). We produced the same result with ABM using Swarm software. Agents were initialized randomly throughout a 200x200 grid. Each agent had a "dialect index" initially set to zero, except agents in a "city" in the center who had a fixed dialect index of 100. Agents then moved randomly for a "time-period" of 120 movements. At the end of each time-period, each agent's dialect index was recalculated by averaging it with the dialect indexes of the other agents it had encountered. After 1,000 time-periods, we plotted all agents' dialect indexes versus their distances from the grid center. The result was a logarithmic curve ($p<0.0000001$), strikingly similar to real-world results. This shows that the model can simulate this fundamental pattern of diffusion predicted by real-world dialectometry, and the pattern emerges from simple density of interactions, not anything specific to human language. (2) Labov (2007) shows that parent-to-child transmission faithfully reproduces structural patterns like the Northern Cities Shift (NCS), but adult-to-adult diffusion does not. NCS is transmitted faithfully to new generations of Inland North children, and then progressively incremented. But St. Louis speakers, depending only on adult-adult contact, only attain an incomplete, unsystematic version. Labov (2007) attributes this difference to children's superior language-learning ability. In our model both adults and children have the same learning mechanism. Simple density and a "blank slate" model of child dialect acquisition are sufficient. Using ABM, we modeled the chain shift with an abstract set of vowel "frequencies" (cf. Ettlinger 2007; Schwartz et al. 1997; Labov 2010:142-4), such as Vowel #1=5.0 units and Vowel #2=15.0 units, etc. When any two agents randomly interact, they adjust their frequencies slightly until they achieve "communication" of a given vowel. They then store each others' vowel frequencies as exemplars, and draw upon those exemplars in subsequent interactions. Crucially, children are initialized as a blank slate (no exemplars). Adults start with 75 exemplars per vowel. We triggered a chain shift in "Chicago" by initializing two vowels close together. The model correctly produced the chain-shift effects observed in real-world Chicago and St. Louis: Child agents acquired shifted vowels from native "Chicago" adults and began incrementing the next stage. The non-shifted "St. Louis" adults in contact with Chicago adults only learned "phonetic variants": Their vowels were influenced by the "Chicago" frequencies, but were not incorporating the full, systematic shift. The results support our hypothesis above.