Language Leaders and Effective Population Size

Robin Clark
Department of Linguistics
University of Pennsylvania

March 23, 2016

In this brief note, I intend to advance a new view on the treatment of Language Leaders in a population. The role of language leaders in sociolinguistic variation has been a topic of enormous interest to the field. Descriptively, language leaders are well-connected individuals at the leading edge of language change; they seem to be innovators who are able to drive variation due to their central position in a social network. I will argue that language leaders are characterized by their ability to replicate features of their idiolect in others; that is, language leaders have enhanced "social reproduction."

To understand the effect leaders can have on a population, consider the Agent-Based Model reported in Clark and Kimbrough (2015). The model consists of 4,000 agents divided into two groups: Lefties and Righties. Each agent has a vowel system consisting of eight vowels (specified by the frequencies of two formants, F1 and F2); these vowels are derived from a queue consisting of one hundred vowels by taking the average of F1 and F2, ($\overline{\text{F1,F2}}$). As a consequence, no two agents have precisely the same representation of the vowel system.

Agents are paired and one agent sends a vowel to the other agent, who must then classify the new vowel. The agent who sends the vowel selects the vowel according to a uniform distribution of the vowels; the "articulatory target" of the agent is determined by ($\overline{\text{F1,F2}}$), although noise is added to the result. The receiver classifies the incoming vowel by taking the ratio of F2 to F1 and comparing it to the ratios of its existing vowel representations. The new instance is then added to the queue of the vowel that they classified

the new vowel as. The oldest vowel on the queue is then removed. Note that the agents receive no feedback as to whether their classification was correct. Clark and Kimbrough (2015) provide a more detailed discussion of the mechanics of the model.

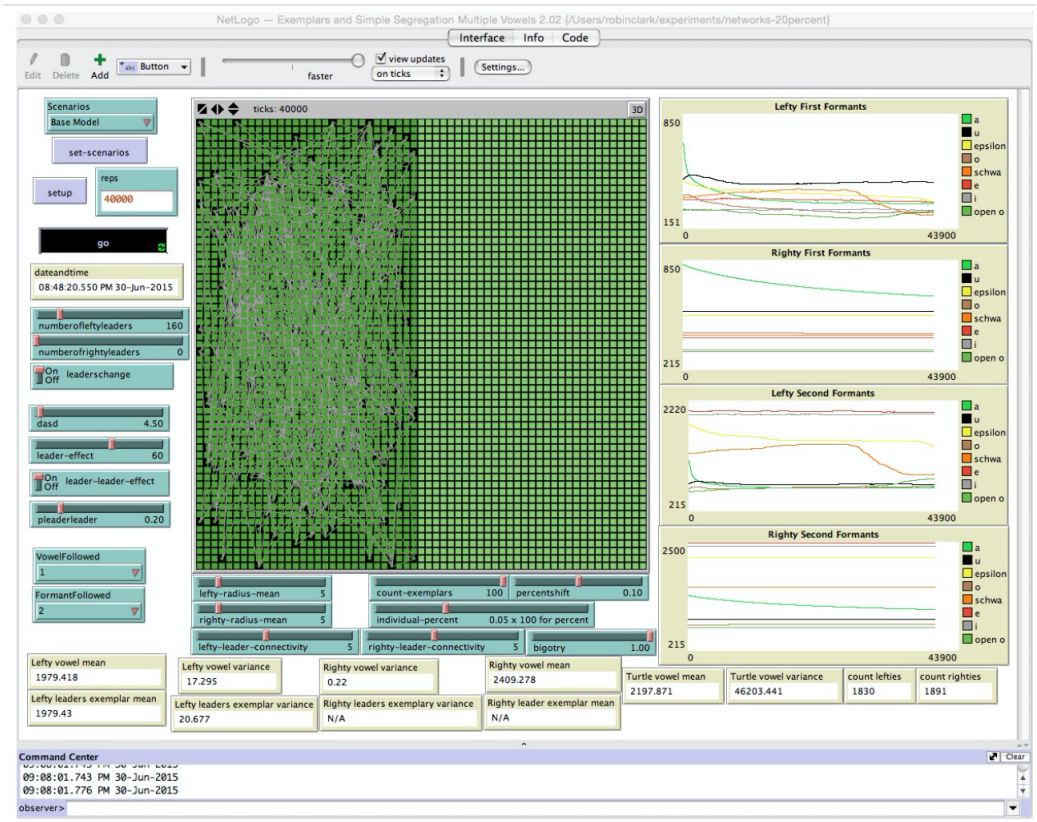

Figure 1: An illustration of the "Leader effect." The vowel system of the Lefties is distorted in the presence of leaders; the Righties, lacking leaders, simply blend.

The main effect can be seen by contrasting the behavior of the Lefties and the Righties with respect to the first and second formants of their vowel systems, shown on the right-hand side of the screengrab in Figure 1. Consider, first, the Righties; the average frequencies for F1 and F2 show that the population is quite stable with no variation, as we would expect. The Righty population is entirely lacking in language leaders. Contrast this with

the Lefties; here, the average frequencies for F1 and F2 show a great deal of instability. The sole difference between the Lefties and the Righties, here, is that the Lefties have 160 language leaders in their population.

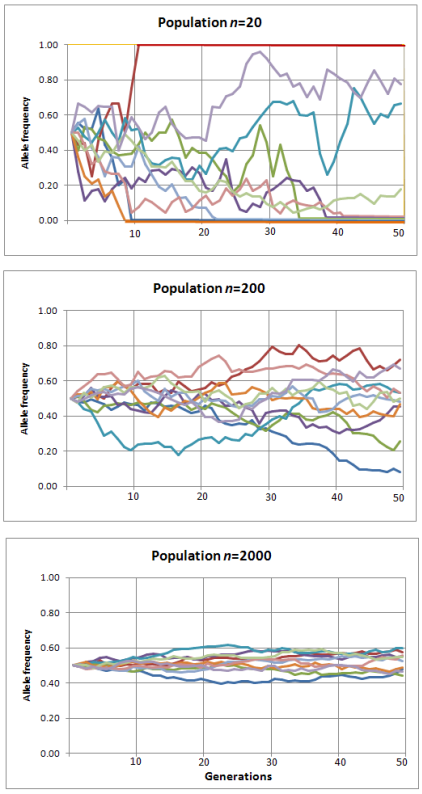

Figure 2: Population size and drift

Our initial analysis was that the subpopulation of language leaders would have a higher variance than the main population and, in addition, these leaders would draw other members of the population in toward them because of their greater influence. We will now propose a more specific analysis of the phenomenon.

Consider the graphs in Figure 2.[1] The figures show examples of drift in allele frequencies in populations of various sizes; notice that variance is greater

[1]From Wikipedia, because I haven't gotten around to generating comparable graphics!

in smaller populations. I propose that we connect the variation observed in the presence of leaders with drift in a more direct fashion. In particular, we can use effective population size in the case where the ratio between males and females are unequal. Crow and Kimura (1970) give the following equation for effective population number:

$$N_e = \frac{4N_m N_f}{N_m + N_f} \tag{1}$$

where N_e is the effective population number, N_m is the number of males and N_f is the number of females. We need to suppose, here, that language leaders have a different impact on the replication of linguistic features than other agents.

Suppose, by way of example, that there are 4,000 agents, 5% of whom are leaders; substituting into the equation we get:

$$N_e = \frac{4 * 200 * 3800}{4000} = 760$$

The equation in 1 is clearly not correct for the case of social transmission and, so, will have to be refined. The point, though, is that the presence of social structure can have an impact on effective population size. In particular, language leaders, because of their enhanced ability to replicate their linguistic forms, introduce the social equivalent of an uneven sex ratio, reducing population size and increasing the level of drift the population could be subject to. This drift could, furthermore, be the source of variant forms which could be picked up as social markers, feeding language variation and change.

References

Clark, R. and Kimbrough, S. O. (2015). The spontaneous emergence of language variation from a homogeneous population the spontaneous emergence of language variation from a homogeneous population. In *https://computationalsocialscience.org/csssa2015-conference-schedule/csssa-2015-papers/*. Computational Social Science Society of the Americas.

Crow, J. F. and Kimura, M. (1970). *An Introduction to Population Genetics Theory*. The Blackburn Press, Caldwell, NJ.