Social networks and intraspeaker variation during periods of language change

Some analysts (Port 2004) have argued that social networks are the primary determinant of linguistic structure. Social networks clearly do play a fundamental role in language change and variation (Milroy & Milroy 1985, Eckert 2000). In particular, Nettle (1999) argues that they solve the THRESHOLD PROBLEM: how is it that rare linguistic variants spread through a speech community (Sapir 1921)? For example, in English, initially rare periphrastic DO (*She does not speak*) replaced the finite verb in negative declaratives (*She speaks not*; Ellegård 1953, Kroch 1989). Previous studies (Nettle 1999) of the threshold problem have focused on language learners that can produce variants consistent with one grammar or the other, but not both (i.e., discrete grammars). However, the assumption that learners are limited to discrete grammars is inconsistent with the presence of intraspeaker variation during language change (Weinreich, Labov & Herzog 1968); e.g. in the same text, individual early English writers produced negative declaratives both with and without periphrastic DO (Ellegård 1953, Kroch 1989). We show that although a social network model (Barabási & Albert 1999) of learners with a specific type of probabilistic language model (Yang 2002) can account for intraspeaker variation, it cannot solve the threshold problem, unless learners have a soft preference for discrete grammars along with a bias towards the minority variant.

In exploring the conditions for the spread of rare variants, we replicated previous findings that show it can occur in social networks composed of learners with discrete grammars. We utilized a preferential attachment-based network structure (Barabási & Albert 1999) where a learner entering a network prefers to interact with learners who already interact with many other speakers. We simulated populations of varying sizes (20, 40, or 60 individuals) 12 times for 12 iterations each. At each iteration, learners adopted the grammar of the majority of their neighbors in the network. Consistent with other studies (Motter & Lai 2002), these networks overcame the threshold problem - the rare variant could in fact spread to become the dominant form.

Can such networks address the threshold problem when learners have probabilistic language models? In a second set of simulations, learners used their interactions with other speakers to assign gradient weights to particular grammatical options using the linear reward/punish algorithm (Bush & Mosteller 1951, Yang 2002). For example, a speaker might learn to produce forms consistent with the periphrastic DO grammar in 80% of utterances. We conducted 12 simulations (8000 iterations each) with varying-sized populations (20, 40, or 60) with a fixed percentage of 25% of learners initially selecting one grammatical option. In every run, the learners' grammars quickly became very similar to one another. However, neither variant came to dominate the population. Instead, learners quickly converged upon some probabilistic norm (e.g., all producing one variant about 30% of the time).

These results show that social networks of learners with discrete grammars can solve the threshold problem but fail to account for intraspeaker variation during language change. In contrast, we find that when learners are equipped with a simple probabilistic language model, these networks are unable to solve the threshold problem. Therefore, we considered an intermediate position: learners equipped with a soft preference for having a single grammatical option (Kroch 1994). We modified the probabilistic model such that each speaker's weighting of their grammatical options was biased towards extreme values. We ran this simulation 1000 times, for 1000 iterations each, limiting our focus to networks of 40 individuals. This time, instead of the population converging to some probabilistic norm, individuals moved toward categorically using one variant or the other. Furthermore, the rate of change of the learners' grammars followed the characteristic S-shaped curve for language change (Labov 1994). However, these networks still could not solve the threshold problem. An initially rare variant was never adopted by the majority of the population (Fig. 1).

To address this issue, we considered a fourth model in which speakers have a soft bias (e.g., based on social or functional factors) towards the initially rare variant. Under the same simulation conditions as the previous model but incorporating an appropriate level of bias, the preferred minority variant could spread through the majority of the network (Fig. 2). Therefore, our results suggest that the conditions for solving the threshold problem while accounting for intraspeaker variation involve a soft preference toward acquiring a single grammatical option coupled with a speaker preference for the minority variant.
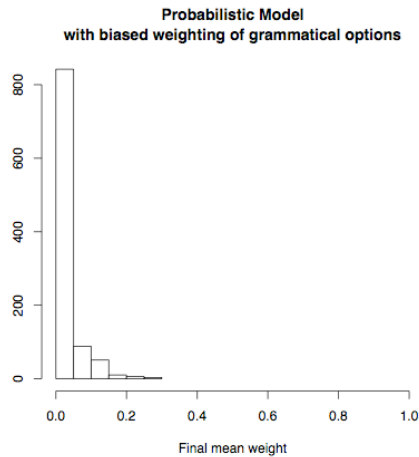
Figure 1: Distribution at the end of 1000 iterations of the mean gradient weight of the initially rare variant for all individuals in the network. When speakers' weightings were biased toward extreme values (ie. toward 0 or 1), the rare variant could not spread.
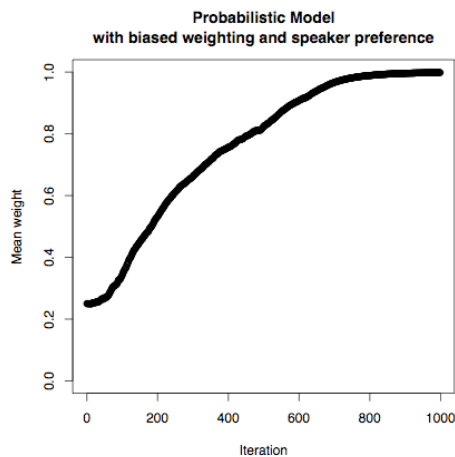

Figure 2: Mean weight over time for one of the 1000 runs in this condition. When incorporating both biased weights and a speaker preference for the initially rare variant, that variant could spread through the entire population, solving the threshold problem.

Selected References
Barabási, A. and Albert, R. (1999). Emergence of scaling in random networks. Science 286: 509-512.
Kroch, A. (1994). Morphosyntactic variation. In K. Beals (ed.). *Proceedings of the thirtieth annual meeting of the Chicago Linguistics Society.* Vol. 2. 180-201.
Motter, A. and Lai, Y-C. (2002). Cascade-based attacks on complex networks. *Physical Review E*, 66, 065102.
Port, R. (2004) Social, Sensory and Symbolic Aspects of Phonology. Ms, Indiana University.
Nettle, D. (1999). Using Social Impact Theory to simulate language change. Lingua 108: 95-117.
Yang, C. (2002). Knowledge and learning in natural language. Oxford: Oxford Univerisity Press.