NATURAL POPULATION GROWTH CAN CAUSE LANGUAGE SIMPLIFICATION

Stella Frank[*1] and Kenny Smith[1]

[*]Corresponding Author: stella.frank@ed.ac.uk
[1]Centre for Language Evolution, University of Edinburgh, UK

Linguistic complexity is known to be negatively correlated with population size. We hypothesise that language simplification could occur during natural population growth as a result of increased numbers of learners in the population. Simulation results confirm this hypothesis: younger age distributions result in lower complexity, independently of population size, and growing populations show a drop in complexity that matches the increase in young learners.

1. Introduction

Larger populations tend to speak simpler languages, as measured by the diversity of morphological and grammatical structures (Lupyan & Dale, 2010). This population size effect has been argued to be a result of population growth via immigration, with adult learners having a simplifying effect on the language (Wray & Grace, 2007; McWhorter, 2007; Trudgill, 2011, see also Bentz & Winter (2013); Bentz et al. (2015) for empirical support). However, in a large-scale analysis, Koplenig (2019) does not find an effect of language "vehicularity" (whether a language is learned by adults) on complexity, but does find an effect of population size. This recent result requires reexamining the proposed mechanism behind the link between population size and linguistic complexity.

An alternate route to a larger population that does not involve immigration (with the attendant adult language learners), is so-called 'natural' population growth as the result of birth rates exceeding death rates. More specifically, if within the population speaking a given language, there are more new speakers (due either to increasing birth rates or decreasing infant and early-age mortality rates), while adult mortality remains constant, the total number of speakers of that language will increase. The demographic consequences of natural population growth is an age distribution that is skewed towards the young, a phenomenon found in contemporary growing populations but also in small-scale societies in the past. Data from modern hunter-gatherer and other small-scale societies indicate that the main source of demographic variation between these groups is the rate of infant and early childhood mortality while adult rates of mortality are relatively similar (Gurven & Kaplan, 2007; Pennington, 1996). Increases in population size, whether enabled by climate or other factors (Tallavaara & Sepp, 2011; Bettinger, 2016), would thus be most likely driven by more children surviving to adulthood, rather than by adults living longer. During the Neolithic transition to agriculture, population demographics changed dramatically in what is called the Neolithic Demographic Transition (Bocquet-Appel, 2011; Shennan, 2001): birth rates increased sharply, and were only later balanced out by increased mortality rates. This led to youth-heavy populations: "At the peak of the NDT, there were children everywhere and the average age of the population was about 18 years old" (Bocquet-Appel, 2011).

What effect could this youth-heavy demographic distribution have on language? Younger speakers still in the process of learning the language are generally simply less accurate speakers of the language. Language changes are also often led by younger speakers, amplified by networks of age-peers learning from and reinforcing one another (Labov, 2007; Cournane, 2017; Sankoff, 2018). A population with more younger speakers, such as a naturally growing population, might thus be subject to more linguistic change, or at least more attempted changes; this could also lead to fewer fully accepted variants, if new idiosyncratic variants displace older, more wide-spread variants in learner's inputs.

In this paper, we use an agent-based population model to demonstrate that the above holds: increasing the proportion of younger speaker agents leads to a drop in the number of linguistic variants shared by every agent in the population, i.e., the simplest measure of linguistic complexity. These shared variants can be thought of as the set of forms that are a required part of a speaker's linguistic inventory. The agents within the model are formulated to be comparable with previous work (Reali et al., 2014, 2018; Spike, 2017) but we introduce more realistic population dynamics, including population turnover.

In our simulations, we first disentangle population size and population demographics by simulating the development of languages in populations with different demographic distributions, e.g. populations in which there are always more younger speakers than older speakers, or in which ages are more evenly distributed, but without population growth. We find that stationary populations with more younger speakers converge to languages with lower complexity, as measured by the number of variants shared throughout the population. Age demographics and size interact, with younger smaller populations having similar complexity levels as older larger populations.

In actually growing populations, in which more agents are added than removed, we then find that language complexity drops as the population increases. The drop in complexity is in proportion to the rate of population growth, i.e., the extent to which the demographic distribution shifts from being evenly distributed to more youth-heavy. This further supports our hypothesis that increasing numbers of younger speakers, due to natural population growth, is an alternate mechanism to language simplification.

2. Model details

2.1. *Agents and their language*

We follow Reali et al. (2014, 2018); Spike (2017) in using a Dirichlet Process or (equivalently) Hoppe Urn model of agent language learning. Agents learn a set of items (corresponding to lexical items or grammatical variants) by interacting with other agents. When they hear an item, it gets added to their lexicon. Agents speak by drawing from the lexicon, where the probability of producing an item is proportional to the number of times (tokens) that item (type) has been stored in the lexicon: $p(x) = \text{Count}(x)/(M+\alpha)$, where M is the number of tokens seen so far and α is an 'innovation' hyperparameter. Namely, agents also have the possibility of inventing a new item, with probability inversely proportional to the number of tokens already in their lexicon: $p_{new} = \alpha/(M + \alpha)$. We set $\alpha = 1$ throughout. Innovated items are always unique (two agents can't separately innovate the same item). Agents also update the counts in their lexicon with their own productions, in order to ensure that their own new innovations are part of their stored lexicon. At the very beginning of the simulation, there are no seeded items: the agents start out with empty lexicons, the same as when agents enter the population later on.

The agents' lexicons may be bounded by a memory limit, in which case only the last m tokens (heard or spoken) are retained in the lexicon and the oldest tokens are deleted. Rare types will thus disappear if they are not used. Note that the memory limit also indirectly sets the lower bound on the probability of generating new items, since M is upper bounded by the memory limit.

Because of their lack of experience, and subsequent smaller M, younger agents will be more influenced more by every interaction they have, and are more likely to adopt new variants, as well as innovate new variants themselves. Older agents on the other hand tend to be more conservative than younger agents, in the sense that when they encounter a new variant, they will add it to their lexicon, but will be unlikely to start producing it themselves. A memory limit will limit the convergence of older agents, and consequently reduce the difference between older and younger agents.

2.2. *Agent interactions*

Agents interact in dialogues in which each agent speaks and then listens to the other agent for 10 turns.[1] Agents participate in 10 dialogues per 'epoch', i.e. between turnover periods, described in the next section.

The population is organised as a fully-connected graph, meaning the probability of any two agents being paired up is equal. As a consequence, young agents in

[1]We found longer dialogues, while not common in previous work, to be important for getting new variants to spread: they need to build up enough probability in the context of first use, in a kind of 'conceptual pact', for the agents to reuse them in subsequent dialogues with other agents.

youth-heavy populations will thus be more likely to speak among themselves and will, as a result, learn language as much (or more) from their age peers than from their elders.

Agents are considered to be adults after an initial learning phase consisting of two epochs, corresponding to 400 updates in 20 dialogues. During this initial learning phase, child agents do not speak, so other agents do not update, analogous to adults ignoring child babbling.

The complexity of a language in a population at a given point in time is measured as the number of variants shared by all adult agents in the population, i.e. the size of the common language. This is the same criterion as used by Spike (2017) and we use it to avoid setting arbitrary thresholds. We have verified that the pattern of results is the same with less stringent criteria, e.g. measuring complexity as the number of variants shared by 50% of the population.

2.3. *Population Turnover*

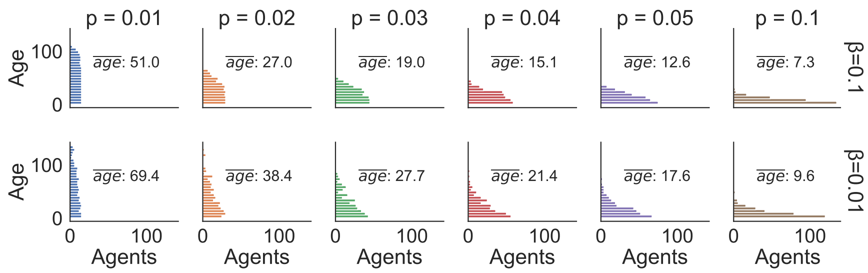

Figure 1. Population pyramids created by removing agents at different turnover rates (columns) using the Gompertz function with two different β settings (rows). Increased turnover leads to younger populations, with a lower mean age.

We implement gradual turnover within our population by selecting a number of agents at each epoch to replace. The turnover parameter p regulates how many agents are removed at each epoch ($p \times N$, the number of agents in the population); in a stable (not growing) population, the same number of agents are replaced with new agents. The more agents are replaced, the more the resulting demographic distributions are weighted towards younger agents. Figure 1 shows these age distributions in the form of 'population pyramids': higher levels of turnover lead to larger numbers of younger agents and a lower average age of the population.

In a realistic population, older agents are more likely to be removed than younger agents. To achieve this dynamic, we use the Gompertz function, which resembles an asymmetric sigmoid and was developed to characterise human life expectancy at higher ages (Gompertz, 1825, see also Baxter & Croft (2016) for use in a model of language change). We follow the parameterisation given in Missov et al. (2015), where the age-dependent probability of picking an agent for removal is $p(age) = \beta \exp \beta(age - M)$, where β is a rate parameter and M is a parameter corresponding to modal age at death (set throughout to 100, although in populations with high turnover, effective age at death will be much younger). Changing the rate parameter β affects the likelihood of picking only the oldest agents: with a smaller setting ($\beta = 0.01$), younger agents are also sometimes removed, leaving some older agents to remain in the population longer.

Importantly, we do not select the agents to remove uniformly at random (cf. Spike, 2017; Dale & Lupyan, 2012; Reali et al., 2018), since this leads to an unrealistic age distribution with mostly young agents but a small number of very long-lived agents. (More formally, age is exponentially distributed as a result of the Poisson point process governing removal.)

2.4. *Related Work*

Agent-based models of language evolution have replicated the effect of population size on language complexity (Dale & Lupyan, 2012; Reali et al., 2014, 2018; Spike, 2017). However, these models either do not include population turnover or do turnover at random, which we argue leads to unrealistic demographic distributions. Similarly, models of sociolinguistic variation and the dynamics of language change (Baxter & Croft, 2016; Stanford & Kenny, 2013; Kauhanen, 2016) have added more realism to network structure and population turnover, but have not investigated the effect of demographic distributions.

In some iterated learning models, populations are modelled as a series of non-overlapping discrete generations, with (vertical) transmission exclusively from older to younger generations (Griffiths & Kalish, 2007; Griffiths & Reali, 2011; Kirby et al., 2014). In contrast, in our population agents enter the population continuously, and learn from interactions. Learning is also symmetrical: within a given interaction, both agents are updating their language, not only the younger agent (though for the reasons spelled out above, the older agent will be influenced less by the younger agent than vice-versa). As a consequence agents can influence other agents that are the same age (horizontal transmission) or older. Another important difference is that, in contrast to 'chain' models in which an agent learns only from a single other (older) cultural parent agent, the agents in our population learn from multiple speakers (see also Niyogi & Berwick, 2009; Smith, 2009; Burkkett & Griffiths, 2009).

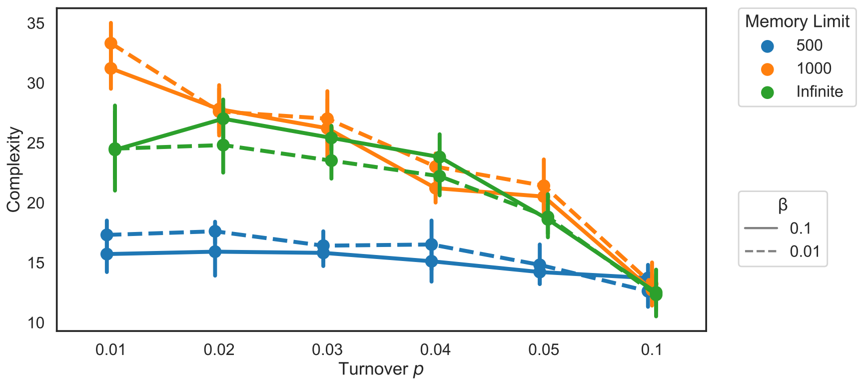

Figure 2. Final complexities of populations with different turnover rates (x- axis), β parameter settings (line styles), and memory limits (colors). Populations become younger going from left to right, in line with a decrease in complexity. Averages shown are over 10 runs.

3. Simulation Results

3.1. *Stable populations with different age distributions*

In our first experiment, we keep the population size stable (at $N = 100$) and vary only the population age distribution by manipulating the rate of turnover. While our broader argument is about population growth driven by an increase in the proportion of young agents, a stable non-growing population allows us to separate the effect of age demographics from population growth. In this setting, any differences in linguistic complexity between populations of the same size are due to the relative numbers of young and older agents.

We run each population simulation for 10000 epochs and report the complexity of the language at the final epoch. The final complexities for the populations in Figure 1 with three different memory limits are show in Figure 2. Higher turnover rates, resulting in younger populations, lead to languages with fewer shared variants. Conversely, older populations have languages with more shared variants, as long as the agents have sufficient memory capacity. With a low memory limit all agents act ‘young’ and complexity is reduced across all turnover rates.

Linguistic complexity thus requires a sufficient proportion of experienced older agents with in the population. The exact shape of the age distribution matters less, since the value of β, which mainly effects the number of older agents, doesn’t have an effect. When the numbers of old and young are not balanced, the young learn from peer interactions rather than from older agents. These populations are characterised by more innovative variants which do not spread through the whole population and thus do not contribute to complexity as measured here.

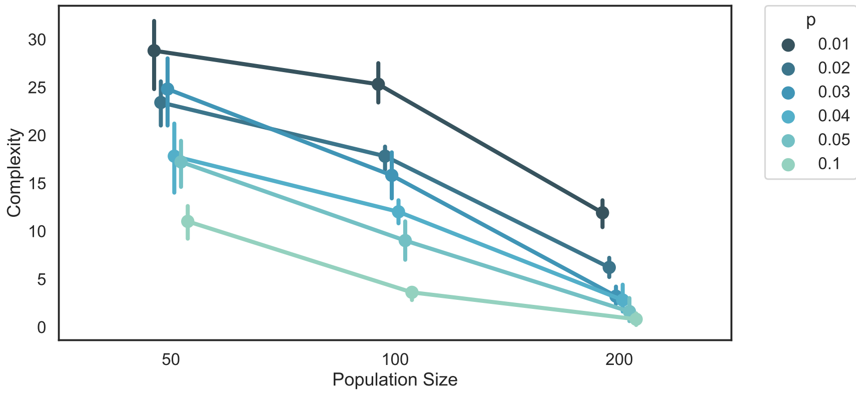

Figure 3. Final complexities of populations of different sizes (x-axis) and demographics (lines). Memory limit $= 1000$, Gompertz $\beta = 0.1$.

3.2. *Population size and age demographics*

The effect that older populations (lower turnover) have higher complexity holds across population sizes (shown in Figure 3, which also shows a replication of the finding that smaller populations have higher complexity across demographics). Intriguingly, smaller youth-heavy populations (i.e. higher p, corresponding to the lower lines on the graph) have similar complexity to larger populations with more evenly distributed age demographic: the former have the demographic signature of growing populations, while the latter are more stable, suggesting a constant level of complexity from a small but growing population to the more stable larger population that is the result of growth.

3.3. *Population growth*

Finally, we check that actual population growth affects the language of the population. In these simulations, we first run a fixed-sized population for long enough for the language to stabilise, at a turnover rate ($p = 0.01$) that results in a population with an even age distribution. During the growth phase, the rate of adding agents is larger than the rate of removing agents by the growth factor g. Depending on g, growth happens faster or slower; we stop growth after the population size has quadrupled.

Population growth leads to an immediate decrease in complexity (Figure 4), with higher rates of growth leading to larger decreases. When growth stops and the population stabilises, complexity increases again, but crucially, at a lower level than before population growth. Natural population growth can thus capture the decrease in complexity linked to larger populations.

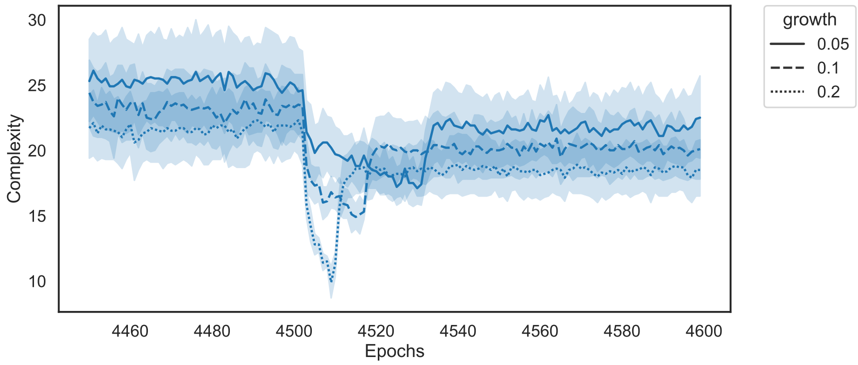

Figure 4. Complexity of a population growing at different rates (g). Growth starts at Epoch 4500 and stops when the population has quadrupled, i.e., $N \geq 400$. Initial $p = 0.01$, no memory limit (infinite), Gompertz $\beta = 0.1$, 10 runs of each setting. Note the initial differences (before Epoch 4500) between g values are due to random fluctuations and are not meaningful.

4. Conclusion

Larger populations can be the result of either migration or natural population growth. Growing populations are characterised by an increased proportion of younger members as compared to stable populations. In our model, populations with these characteristics achieve lower complexity in their linguistic system than populations with the even age distribution associated with stable populations. This offers an alternative route to explaining the link between population size and language complexity that does not involve non-native speakers.

Our results are a consequence of the different learning environments of agents in youth-heavy vs. stable, older, populations. In stable populations, the youngest agents will interact mostly with older agents from whom they learn the language shared by the rest of the population. In contrast, in youth-heavy populations, young agents interact more with their age peers, who have similarly not been exposed to the full language, and thus complexity is lost. The fully-connected network assumed by our model is unrealistic for human populations which have more differentiated social networks. However, in small-scale networks, fully connected networks have similar characteristics to more realistic small-world networks (Spike, 2017). In exploratory modelling experiments, we found that early learner network connectivity (e.g. learning from a principal caretaker) could not mitigate the demographic effect presented in this paper. Secondly, the argument here in a relative one: are children in youth-heavy populations interacting more with, and learning more from, their peers than children in older populations? The answer to this is likely to be yes, though will depend on culturally-specific child-rearing practices and institutions.

References

Baxter, G., & Croft, W. (2016). Modeling language change across the lifespan: Individual trajectories in community change. *Language Variation and Change*, *28*(02), 129–173.

Bentz, C., Verkerk, A., Kiela, D., Hill, F., & Buttery, P. (2015). Adaptive communication: Languages with more non-native speakers tend to have fewer word forms. *PLOS ONE*, *10*(6), e0128254.

Bentz, C., & Winter, B. (2013). Languages with more second language learners tend to lose nominal case. *Language Dynamics and Change*, *3*, 1–27.

Bettinger, R. L. (2016). Prehistoric hunter–gatherer population growth rates rival those of agriculturalists. *Proceedings of the National Academy of Sciences*, *113*(4), 812–814.

Bocquet-Appel, J.-P. (2011). When the world's population took off: The springboard of the Neolithic demographic transition. *Science*, *333*(6042), 560–561.

Burkkett, D., & Griffiths, T. L. (2009). Iterated learning of multiple languages from multiple teachers. In *Proceedings of Evolang.*

Cournane, A. (2017). In defense of the child innovator. In Éric Mathieu & R. Trueswell (Eds.), *Micro-change and macro-change in diachronic syntax* (chap. In defence of the child innovator). Oxford University Press.

Dale, R., & Lupyan, G. (2012). Understanding the origins of morphological diversity: The linguistic niche hypothesis. *Advances in Complex Systems*, *15*(03n04), 1150017.

Gompertz, B. (1825). On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies. *Philosophical Transactions of the Royal Society of London*, *115*(0), 513–583.

Griffiths, T. L., & Kalish, M. L. (2007). Language evolution by iterated learning with Bayesian agents. *Cognitive Science*, *31*(3), 441–480.

Griffiths, T. L., & Reali, F. (2011). Modelling minds as well as populations. *Proceedings of the Royal Society B: Biological Sciences*, *278*(1713), 1773–1776.

Gurven, M., & Kaplan, H. (2007). Longevity among hunter-gatherers: A cross-cultural examination. *Population and Development Review*, *33*(2), 321–365.

Kauhanen, H. (2016). Neutral change. *Journal of Linguistics*, *53*(02), 327–358.

Kirby, S., Griffiths, T., & Smith, K. (2014). Iterated learning and the evolution of language. *Current Opinion in Neurobiology*, *28*, 108–114.

Koplenig, A. (2019). Language structure is influenced by the number of speakers but seemingly not by the proportion of non-native speakers. *Royal Society Open Science*, *6*(2), 181274.

Labov, W. (2007). Transmission and diffusion. *Language*, *83*(2), 344–387.

Lupyan, G., & Dale, R. (2010). Language structure is partly determined by social structure. *PLoS ONE*, *5*(1), e8559.

McWhorter, J. (2007). *Language interrupted: Signs of non-native acquisition in standard language grammars*. Oxford University Press.

Missov, T. I., Lenart, A., Nemeth, L., Canudas-Romo, V., & Vaupel, J. W. (2015). The Gompertz force of mortality in terms of the modal age at death. *Demographic Research*, *32*, 1031–1048.

Niyogi, P., & Berwick, R. C. (2009). The proper treatment of language acquisition and change in a population setting. *Proceedings of the National Academy of Sciences*, *106*(25), 10124–10129.

Pennington, R. L. (1996). Causes of early human population growth. *American journal of physical anthropology*, *99*, 259–274.

Reali, F., Chater, N., & Christiansen, M. H. (2014). The paradox of linguistic complexity and community size. In *The evolution of language: Proceeedings of the 10th international conference.*

Reali, F., Chater, N., & Christiansen, M. H. (2018). Simpler grammar, larger vocabulary: How population size affects language. *Proceedings of the Royal Society B: Biological Sciences*, *285*(1871), 20172586.

Sankoff, G. (2018). Language change across the lifespan. *Annual Review of Linguistics*, *4*(1), 297–316.

Shennan, S. (2001). Demography and cultural innovation: a model and its implications for the emergence of modern human culture. *Cambridge Archaeological Journal*, *11*(01).

Smith, K. (2009). Iterated learning in populations of Bayesian agents. In *Proceedings of CogSci.*

Spike, M. (2017). Population size, learning, and innovation determine linguistic complexity. In *Proceedings of CogSci.*

Stanford, J. N., & Kenny, L. A. (2013). Revisiting transmission and diffusion: An agent-based model of vowel chain shifts across large communities. *Language Variation and Change*, *25*(02), 119–153.

Tallavaara, M., & Sepp, H. (2011). Did the mid-holocene environmental changes cause the boom and bust of hunter-gatherer population size in eastern fennoscandia? *The Holocene*, *22*(2), 215–225.

Trudgill, P. (2011). *Sociolinguistic typology: Social determinants of linguistic complexity*. Oxford University Press.

Wray, A., & Grace, G. W. (2007). The consequences of talking to strangers: Evolutionary corollaries of socio-cultural influences on linguistic form. *Lingua*, *117*(3), 543–578.