LINGUISTIC STABILITY INCREASES WITH POPULATION SIZE, BUT ONLY IN STABLE LEARNING ENVIRONMENTS

ANDREAS BAUMANN[*1]

[*]Corresponding Author: andreas.baumann@univie.ac.at

[1]Department of English and American Studies, University of Vienna, Vienna, Austria

The effect of population size on linguistic stability and evolution has been investigated in different linguistic domains. The relationship among these factors, however, is not always clear. In this paper, we study a basic population-dynamical model of linguistic spread, derive measures of linguistic stability and fitness, and investigate the effect of population size on these measures. By allowing for stochasticity in the learning process of linguistic constituents, it is shown that a constituent's stability and fitness increases with population size, but that high variability in the learning environment may cause constituent loss, also in large populations. The respective roles of learning and usability are also discussed.

1. Population size and linguistic evolution

Population size has been proposed to affect linguistic structure (Atkinson, Kirby, & Smith, 2015; Hay & Bauer, 2007; Lupyan & Dale, 2010; Nettle, 2012; Wichmann, Rama, & Holman, 2011) as well as rate of linguistic change (Atkinson, 2011; Wichmann & Holman, 2009) and degree of adaptation with respect to cognitive and communicative pressures (Fay & Ellison, 2013). More recently, Bromham et al. (2015) have shown in their empirical study that lexical items are more stable in large populations and that rates of word loss are higher in small populations. Indeed, if linguistic constituents share mechanistic similarities with biological replicators (Croft, 2000; Ritt, 2004) the latter observation is exactly what one would expect as per evolutionary theory (Bromham et al. 2015: 2100).

Purely computational approaches to this problem have been, to our knowledge, primarily limited to simulations (Nettle, 1999; Wichmann, Stauffer, Schulze, & Holman, 2008). More recent advances in mathematical ecology (in particular, stochastic epidemiological dynamics; Gray, Greenhalgh, Hu, Mao, & Pan, 2011; Greenhalgh, Liang, & Mao, 2015) allow for a more analytical assessment. This paper adds to the discussion about the relationship between

population size, linguistic stability and evolution by modifying and analyzing an established population-dynamical model of linguistic spread (Cavalli-Sforza & Feldman, 1981; Nowak, 2000; Nowak, Plotkin, & Jansen, 2000; Solé, Corominas-Murtra, & Fortuny, 2010; Wang & Minett, 2005). We focus on the dynamics of single 'linguistic items' or 'constituents' (like phonemes, n-phones, words or constructions) in finite speaker populations. After discussing the deterministic dynamics, we also analyze a stochastic version of the model, which accounts for variability in the process of constituent learning (e.g. varying density of the speaker network due to eco-linguistic factors, or varying usage of the constituent in learner-user interactions). It is shown that the general assumption that linguistic stability increases with population size only holds if variability in the learning process is kept low, and argue that the latter factor provides an interesting mechanism in language evolution.

2. Modeling linguistic spread in finite populations

2.1. Deterministic model

We study a modified version of Nowak's (2000) basic model of linguistic spread. In our version of the model, population size N is restricted to be finite. The model describes the dynamics of a structured population composed of users of a particular linguistic item i (e.g. phoneme, n-phone, word or construction) and learners that do not use it. Let U_i and L_i denote the respective sizes of the (disjoint) subpopulations and let $U_i + L_i = N$. Whenever learners and users meet, the former learn i at a rate λ so that they switch from class L_i to U_i. We assume λ to denote the *learning* rate, where learning of a new form is not necessarily restricted to the first years of language acquisition. Rather, we mean any interaction of individuals one of which does not yet know and use a given item. In Nowak's (2000) model, this rate λ is a function of (a) *network density*, linked to the number of communicative encounters a learner is exposed to, (b) *production rate*, i.e. the extent to which the item is produced, and (c) *learnability*, i.e. the probability that the item is successfully acquired when a learner is exposed to it. Learners and users die at a normalized mortality rate of 1 (so that each time unit equals one speaker generation), and dead learners and users are immediately replaced by new individuals that are added to the learner class so that population size is kept constant. In addition, users can switch back to class L_i at a rate γ when they stop using i ('unlearning'), for instance because they forget the item or because they abandon it in favor of a competing linguistic variant. We suggest that γ is inversely related with the *usability* of i in everyday speech events in

which no user-learner interactions are involved. Table 1 summarizes the model parameters.

Table 1. Variables in the model and how they can be interpreted

Variable	Linguistic and cognitive interpretation
N	Total size of the population of linguistic agents composed of *U users* and *L learners*
λ	Item-specific *learning rate* in interactions; depends on *network connectivity* (linked to number of communicative encounters), *production rate* (linked to utterance frequency and ease of production), and *learnability* (linked to ease of perception)
γ	Rate at which individuals stop using an item (in addition to speaker death; rate of 'unlearning'); inversely related to factors enhancing *usability* (e.g. ease of memorization or ease of production); assumed to be independent from learner-user interactions
R_0	Expected number of learners that successfully learn an innovation from a single user

The dynamics are determined by a deterministic two-dimensional dynamical system in continuous time which models the respective growth rates of L_i to U_i. In what follows we will omit the index i, for the sake of simplicity, since we only focus on the dynamics of a single item (although the parallel evolution of several items clearly can be studied as well). The model equations read:

$$
\begin{aligned}
\mathrm{d}L/\mathrm{d}t &= \overbrace{-\lambda LU}^{\text{learning}} + \overbrace{\gamma U}^{\text{unlearning}} - \overbrace{L}^{\text{death}} + \overbrace{N}^{\text{birth}} \\
\mathrm{d}U/\mathrm{d}t &= \underbrace{\lambda LU}_{\text{learning}} - \underbrace{(1+\gamma)U}_{\text{death and unlearning}}
\end{aligned}
\tag{1}
$$

If $\gamma = 0$ and $N = 1$ the dynamical system reduces to the model of linguistic spread in Nowak (2000) and Solé (2011), which is equivalent with a one-dimensional model of logistic growth (although the dynamics can be modeled by a single equation, e.g. only the second one in (1), we stick to the more explicit definition for the sake of clarity).

The qualitative behavior of the model can be predicted by the basic reproductive ratio R_0 which is defined as the expected number of learners that learn an item which has been innovatively introduced into the population by a single user (cf. Nowak 2000, Heffernan, Smith, & Wahl, 2005). If $R_0 > 1$ the dynamics approach a non-trivial equilibrium so that $\widehat{U} = N(1 - 1/R_0)$ users know and use the item. That is, the item is stably established in the linguistic community. If, however, $R_0 < 1$ then the dynamics approach an equilibrium in which $\widehat{U}_0 = 0$ users know the item. In that case, the item drops out of usage. Thus,

the basic reproductive ratio functions as a measure of the stability of a linguistic item. For that reason we treat R_0 as a measure of diachronic stability (technically, R_0 measures the stability of the equilibrium $\hat{U}_0 = 0$; if $R_0 > 1$ then $\hat{U}_0$ is unstable so that the population of users persists with probability 1 if any users are added to the population; if $R_0 > 1$ then $\hat{U}_0$ is stable so that the population of users goes extinct with probability 1).

For the present model, the basic reproductive ratio can be shown to read $R_0 = N\lambda/(1 + \gamma)$. The formula can be intuitively understood in the following way. The expected time an individual knowing the item remains in the user class is $1/(1 + \gamma)$; based on our assumption that the item is an innovation there are (approximately) N individuals that do not yet know the item; and each learner acquires the item at a rate of λ. Note, crucially, that since the amount of individuals which can acquire an item from a user depends on the number of learners available in the population R_0 depends on population size. Here this dependency is linear, which is an immediate reflex of the assumption that the population is homogeneously mixed so that any user can inform any learner in the population (see Section 3 for some discussion).

We are interested in the role that population size plays for the stability of a linguistic item. The basic reproductive ratio R_0 increases with N since $\lambda/(1 + \gamma) > 0$. The larger the population, the less likely is it that R_0 falls below one so that the item would inevitably drop out of usage.

In evolutionary terms, $R_0(\lambda, \gamma)$ can be interpreted as a measure the fitness of a linguistic item (Metz, Mylius, & Diekmann, 1996). Evidently, R_0 increases with λ (because $\partial R_0/\partial\lambda = N/(1 + \gamma) > 0$) and decreases with γ (because $\partial R_0/\partial\lambda = -N\lambda/(1 + \gamma)^2 < 0$). Thus, items with high learning rates and high usability should be selected for. That is, items are expected to evolve in such a way that they maximize ease of acquisition, production and use (probably governed by some trade-off among these factors). Moreover, the effect of optimizing λ and γ gets stronger the larger the population size N, so that items are expected to be less optimized in small populations.

2.2. Stochastic model

Things get slightly more complicated when variability in the model dynamics is considered. For instance, demographic variability could be accounted for, i.e. fluctuations due to random speaker deaths and births in addition to the deterministic model dynamics. For the class of models (1) belongs to, it has been shown that the effects of demographic variability can be neglected if population

size is substantially large (Greenhalgh et al. 2015).[1] Another source of variability might be more relevant to linguistic dynamics, namely that of parametric (or environmental) variability. Here, model parameters fluctuate randomly, thus affecting the behavior of all individuals in the population at the same time. In a linguistic setting, for instance, network density of the entire speech community could vary due to eco-linguistic factors (e.g. migration or areal expansion; cf. Mufwene, 2001; Lupyan & Dale, 2010). Likewise, frequency of use of an item established in a speech community might fluctuate due to socio-linguistic or language-internal factors (e.g. morpho-syntactic or phonological restructuring, or emergence of competing variants for instance in language contact). All of these factors can be argued to have an impact on the linguistic learning process. Thus, we include a stochastic component into the model by extending the rate of transition from class L to class U, denoted by $\tilde{\lambda}$, so that $\tilde{\lambda}dt = \lambda dt + \sigma dW(t)$. Here, $W(t)$ is a Wiener process (random noise) which accounts for fluctuation around λ, and $\sigma \geq 0$ is the variance in the 'learning environment' due to the above-mentioned factors. Thus, σ measures the magnitude of these fluctuations. We consider learning environments with low σ as more stable that those with large σ.[2] By replacing λdt by $\tilde{\lambda}dt$ in (1), the model becomes a system of stochastic differential equations (SDE; Allen, 2010):

$$
\begin{aligned}
dL &= (-\lambda LU + \gamma U - L + N)dt - \sigma LU dW(t) \\
dU &= (\lambda LU - (1+\gamma)U)dt + \sigma LU dW(t)
\end{aligned}
\tag{2}
$$

Clearly, if there is no fluctuation ($\sigma = 0$), (2) reduces to the deterministic model (1). System (2) belongs to the class of Itô SDEs analyzed by Gray et al. (2011). Hence, we can employ the conditions for extinction and persistence derived there. By applying Theorem 4.1 in Gray et al. (2011), the basic reproductive ratio for system (2) can be shown to read

$$
R_0 = \underbrace{\frac{\lambda N}{1+\gamma}}_{(i)} - \underbrace{\frac{\frac{1}{2}\sigma^2 N^2}{1+\gamma}}_{(ii)}
\tag{3}
$$

where part (i) equals the basic reproductive ratio of the deterministic system (1) and part (ii) comes from the diffusion term in the SDE (2). Theorem 5.1 in Gray et al. (2011) entails that the system leads to persistence of an item (i.e. stable and positive U), if $R_0 > 1$. If, on the contrary, $R_0 < 1$ and $\sigma \leq \sqrt{\lambda/N}$ (Thm 4.1), or

---

[1] Based on Greenhalgh et al. (2015, Theorem 4.1), demographic variability only has an additional effect if population size falls below critical size $N_{\text{crit}} = 1/4 + (1+\gamma)/\lambda$.

[2] Note that this notion of stability differs from the one measured by the basic reproductive ratio. While σ measures how constantly transmission of an item takes place, R_0 measures whether or not an item persists in the speaker population.

$\sigma > \sqrt{\lambda/N}$ (Thm 4.3), then the number of users U approaches zero with probability 1, so that the item goes extinct.

Several observations can be made. To begin with, it is not difficult to see that $dR_0/dN > 0$ if $\sqrt{\lambda/N} > \sigma$. This means that the stability of an item increases with population size N as long as variability is not too high. In particular, inequality $\sigma > \sqrt{\lambda/N}$ is favored to hold (a) if learning variability σ is large or (b) if population size is high (or both). Thus, severe fluctuations promote the loss of items and impede the establishment of new items in the speaker population (Figure 1). Moreover, for fixed σ, larger population sizes can also have negative effects on the stability of linguistic items. In large populations, even mild fluctuations can yield severe reflexes, as long as they affect the entire linguistic population.
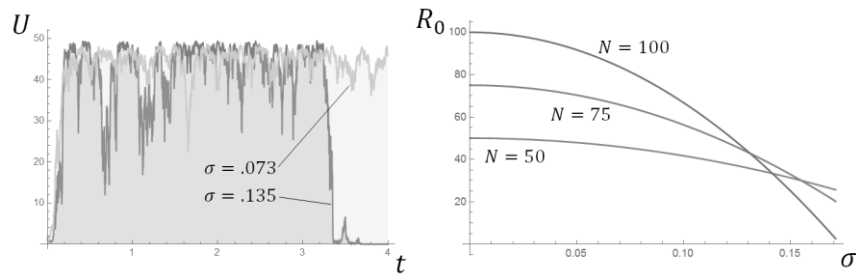

Figure 1. On the left: Itô-process simulations of diachronic developments ($N = 50, \lambda = 0.5, \gamma = 1.5, U(0) = 1$) in two different environments; lower variability ($\sigma = .073$, $R_0 = 7.33$, light gray), and higher variability ($\sigma = .135$, $R_0 = 0.89$, dark gray). After about 3 generations, the item exposed to higher variability in the learning environment goes extinct, as expected. On the right: R_0 as a decreasing function of σ for three different population sizes $N = 50; 75; 100$ ($\lambda = 0.5, \gamma = 1.5$ fixed). For high σ, larger populations yield lower R_0. Computations were done in Mathematica (Wolfram Research, 2016).

What is more interesting is this: a sensitivity analysis reveals information about the relative importance of λ and γ in the optimization of R_0 in the stochastic model. For the respective directional derivatives of $R_0(\lambda, \gamma)$, we have that $\partial R_0(\lambda, \gamma)/\partial(1,0) = N/(1 + \gamma) > 0$, and that $\partial R_0(\lambda, \gamma)/\partial(0, -1) = 1/2 \cdot N(2\lambda - \sigma^2 N^2) \log(1 + \gamma) > 0$, because $\sqrt{\lambda/N} > \sigma$ if the item already exists stably. Items benefit from increasing λ and decreasing γ (i.e. increasing usability), but in contrast to the former parameter, the effect of decreasing γ suffers from variability in the learning environment. For an item, to put it casually, it pays off to put more effort into improving learning rather than usability if variability is high enough. Improving factors that determine learning does always contribute to an item's success, while effects of increased usability may be

vanishingly small in the presence of noise. As in the deterministic case, the effects of optimizing R_0 (i.e. the directional derivatives shown above) get stronger the larger the speaker population (Figure 2).
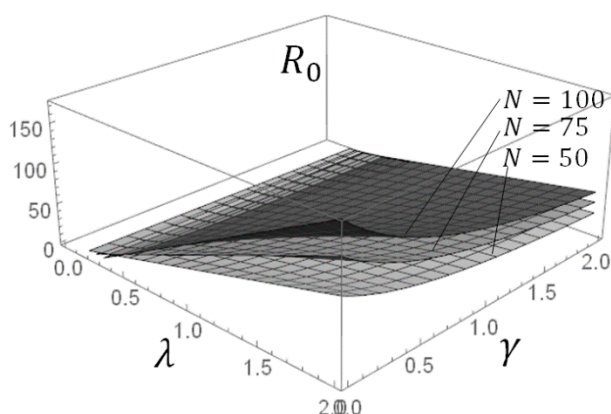

Figure 2. Fitness landscape defined by R_0 as a function of γ and λ for three different population sizes ($N = 50; 75; 100$) in the presence of learning variability ($\sigma = .073$); fitness increases linearly with λ and decreases convexly with γ. Directional slopes (effects of changing parameters) get steeper as population size increases.

3. Discussion and conclusion

By studying systems of ODEs and SDEs, we have shown that population size in general increases the stability of linguistic constituents (cf. Table 2) and thus (a) facilitates their establishment in the speaker population and (b) prevents their loss. This goes in line with Bromham et al. (2015: 2100) who show that "Polynesian languages with larger speaker-population sizes [have] higher rates of gain of new words than their smaller sister languages" and that "languages with a smaller number of speakers [have] higher rates of loss of lexemes". Our results also converge with studies that found a positive correlation between population size and the size of a language's phoneme inventory (see Nettle, 2012 for a review), and by implication phonotactic richness (Maddieson, 2013).

However, the presence of variability in the learning environment decreases stability, and the negative effects of variability get stronger, the larger population size. In the extreme case, this variability can lead to the loss of a constituent (Figure 1, left, dark gray trajectory). As a corollary of this, we can conclude that the establishment of an inventory of constituents (e.g. lexicon of words or

phoneme inventory) requires a relatively stable learning environment (cf. McMahon & McMahon, 2013, p. 248). For instance, it can be argued that the small phoneme inventories found in a number of remote Polynesian languages (Trudgill, 2004) might be a reflex of migration and concomitant variability in network density. We argue that complementary to demographic variability (linked to linguistic founder effects as suggested by Atkinson 2011), environmental variability provides another interesting mechanism for explaining linguistic evolution, because it applies even if population size remains constant.[3]

Table 2. Results and model comparison

Feature	Deterministic model	Stochastic model
Learning environment	Constant ($\sigma = 0$)	Variable ($\sigma > 0$)
Effect of N on stability of constituent	Stability increases with population size N	Stability increases with N if variability σ is small
Effect of N on evolution of learning rate λ	Adaptive effects of improving learning increase with N	Adaptive effects of improving learning increase with N
Effect of N on evolution of usability $\sim \gamma^{-1}$	Adaptive effects of improving usability increase with N	Adaptive effects of improving usability increase with N
Effect of variability on evolution of learning rate λ	Improving learning rate always increases fitness	Improving learning rate always increases fitness
Effect of variability on evolution of usability $\sim \gamma^{-1}$	Improving usability always increases fitness	Effects of improving usability are mitigated by variability σ

One might wonder, what the prediction of the model, that constituent inventories are more likely to shrink in small populations actually means. Clearly, it is not plausible that small populations simply drop constituents like phonemes or lexemes, since some items obviously fulfil specific functions in the linguistic system and cannot be arbitrarily left away. Models like the ones studied in this paper cannot easily account for such details. However, one way of looking at this prediction is this: if constituents vanish (e.g. due to bad adaptation) the language must compensate for this loss, e.g. by adding more complex morpho-syntactic rules. Indeed, this is supported by Lupyan and Dale (2010) who show that small

---

[3] Indeed, Bybee (2011) has contested demographic variability as the main explanatory link between linguistic evolution and population size.

populations sustain morphologically more complex languages.[4] This argument contrasts the causal directionality proposed by Nettle (2012) who argues that it is the smaller number of contacts in small populations that promotes the acquisition of complex morphology (which, in turn, would allow for a reduced lexicon).

The findings also agree with Fay and Ellison (2013: 7) in the sense that increased population size enhances the optimization of properties associated with linguistic transmission. That is, evolution proceeds faster in large populations. At first sight, this may seem paradox: population size is predicted to increase the stability of an item, but at the same time population size drives linguistic optimization, where an item is effectively replaced by a more successful version of itself. Note, crucially, that the more optimized variant is less likely to get lost.

The analysis of the stochastic model has revealed that constituents always benefit from optimizing factors related to learning while advantages gained from optimizing factors related to usability can be lost due to random fluctuations in the learning environment. Based on this, it can be expected that items are relatively more optimized for being learned easily rather than for ease of use outside of the learning context. This accords with studies that propose a strong connection between (diachronic) stability and ease of acquisition (e.g. Monaghan 2014). It is less compatible with studies stressing the importance of usability and ease of production (i.e. speaker-over-listener dominance) in linguistic transmission (Bybee, 2010; Fay & Ellison, 2013).[5]

Finally, a more technical caveat is in order. The model builds on the assumption that the learning process depends on a mass-action law (i.e., interactions are proportional with the product of the number of learners and users). It has been pointed out (de Jong, Diekmann, & Heesterbeek, 1996), that this assumption does not hold in large populations in realistic ecological scenarios. Consequently, the effect of population size on the basic reproductive ratio is probably overestimated as populations become larger. Accounting for these issues

---

[4] Note that this observation does not directly follow from the present analysis but rather represents a tentative hypothesis which is compatible with our results. It would be interesting, however, to study a model which includes the possibility of combining items (perhaps similar to the approach adopted by Nowak et al. 2000) to account for complexity. We would like to thank an anonymous reviewer for pointing this out.

[5] This observation, however, might be grounded in the abstract and simplified way in which learning and using constituents is built into the model. Arguably, the rough distinction between factors relevant to learning interactions and those not associated with interactions is very simplistic and must be refined in order to capture learnability and usability more accurately.

eventually requires the implementation of a more complicated network structure.[6] The observations made in this contribution, nevertheless, do not contradict with results from network epidemiology. In large networks, the invasion threshold vanishes under the assumption of a more realistic network structure (small world; scale free). As a consequence of the presence of super spreaders, items can spread easily through large populations (Barabási, 2016). The effects of fluctuations during the learning process in more realistic networks, though, is yet to be looked at more closely.

Acknowledgements

References

Allen, L. J. S. (2010). *An introduction to stochastic processes with applications to biology*. Boca Raton, Florida: CRC Press.

Atkinson, M., Kirby, S., & Smith, K. (2015). Speaker Input Variability Does Not Explain Why Larger Populations Have Simpler Languages. *PloS one*, *10*(6), e0129463. https://doi.org/10.1371/journal.pone.0129463

Atkinson, Q. D. (2011). Phonemic diversity supports a serial founder effect model of language expansion from Africa. *Science*, *322*(6027), 346–349.

Barabási, A.-L. (2016). *Network science*. Cambridge: Cambridge University Press.

Bromham, L., Hua, X., Fitzpatrick, T., & Greenhill, S. J. (2015). Rate of language evolution is affected by population size. *Proceedings of the National Academy of Sciences*, *112*(7), 2097–2102.

Bybee, J. (2010). *Language, Usage and Cognition*. Cambridge: Cambridge University Press.

Bybee, J. (2011). How plausible is the hypothesis that population size and dispersal are related to phoneme inventory size? Introducing and commenting on a debate. *Linguistic Typology*, *15*, 147–153. https://doi.org/10.1515/LITY.2011.009

---

[6] As one of the reviewers of this contribution has rightfully pointed out, reproductive success is – at least partially – a function of network structure, which in turn determines the number of interactions. Here, the functional relationship between population size and network structure is crucial.

Cavalli-Sforza, L. L., & Feldman, M. W. (1981). *Cultural transmission and evolution: aquantitative approach*. Princeton: Princeton University Press.

Croft, W. (2000). *Explaining language change: An evolutionary approach. Longman linguistics library*. Harlow, England, New York: Longman.

Fay, N., & Ellison, T. M. (2013). The cultural evolution of human communication systems in different sized populations: usability trumps learnability. *PloS one*, *8*. https://doi.org/10.1371/journal.pone.0071781

Gray, A., Greenhalgh, D., Hu, L., Mao, X., & Pan, J. (2011). A stochastic differential equation SIS epidemic model. *SIAM Journal on Applied Mathematics*, *71*(3), 876–902.

Greenhalgh, D., Liang, Y., & Mao, X. (2015). Demographic stochasticity in the SDE SIS epidemic model. *Discrete and Continuous Dynamical Systems-Series B*, *20*(9), 2859–2884.

Hay, J., & Bauer, L. (2007). Phoneme inventory size and population size. *Language*, *83*(2), 388–400.

Heffernan, J., Smith, R., & Wahl, L. (2005). Perspectives on the basic reproductive ratio. *Journal of The Royal Society Interface*, *2*(4), 281–293. https://doi.org/10.1098/rsif.2005.0042

Jong, M. de, Diekmann, O., & Heesterbeek, J. A. P. (1996). How does transmission of infection depend on population size? In *Epidemic models: their structure and relation to data* (pp. 84–94). Cambridge, UK: Cambridge University Press.

Lupyan, G., & Dale, R. (2010). Language structure is partly determined by social structure. *PloS one*, *5*. https://doi.org/10.1371/journal.pone.0008559

Maddieson, I. (2013). Syllable Structure. In M. S. Dryer & M. Haspelmath (Eds.), *The World Atlas of Language Structures Online*. Max Planck Digital Library. Retrieved from http://wals.info/chapter/12

McMahon, A. M. S., & McMahon, R. (2013). *Evolutionary linguistics. Cambridge textbooks in linguistics*. Cambridge [England], New York: Cambridge University Press.

Metz, J.A.J., Mylius, S. D., & Diekmann, O. (1996). When Does Evolution Optimize? On the Relation Between Types of Density Dependence and Evolutionarily Stable Life History Parameters. *IIASA Working Papers*, *96*(004).

Mufwene, S. S. (2001). *The ecology of language evolution*. Cambridge: Cambridge University Press.

Nettle, D. (1999). Is the rate of linguistic change constant? *Lingua*, *108*(2), 119–136.

Nettle, D. (2012). Social scale and structural complexity in human languages. *Phil. Trans. R. Soc. B*, *367*, 1829–1836.

Nowak, M. A. (2000). The basic reproductive ratio of a word, the maximum size of a lexicon. *Journal of theoretical biology*, *204*(2), 179–189. https://doi.org/10.1006/jtbi.2000.1085

Nowak, M. A., Plotkin, J., & Jansen, V. (2000). The evolution of syntactic communication. *Nature*, *404*(6777), 495–498.

Ritt, N. (2004). *Selfish sounds and linguistic evolution: A Darwinian approach to language change*. Cambridge: Cambridge University Press.

Solé, R. V., Corominas-Murtra, B., & Fortuny, J. (2010). Diversity, competition, extinction: the ecophysics of language change. *Journal of The Royal Society Interface*, *7*(53), 1647–1664. Retrieved from http://groups.lis.illinois.edu/amag/langev/paper/sole2010diversitygsc.html

Trudgill, P. (2004). Linguistic and social typology: The Austronesian migrations and phoneme inventories. *Linguistic Typology*, *8*(3), 305–320.

Wang, W., & Minett, J. (2005). The invasion of language: Emergence, change and death. *Trends in Ecology and Evolution*, *20*, 263–269.

Wichmann, S., & Holman, E. (2009). Population size and rates of language change. *Human Biology*, *81*(2), 259–274.

Wichmann, S., Rama, T., & Holman, E. (2011). Phonological diversity, word length, and population sizes across languages: The ASJP evidence. *Linguistic Typology*, *15*(2), 177–197.

Wichmann, S., Stauffer, D., Schulze, C., & Holman, E. (2008). Do language change rates depend on population size? *Advances in Complex Systems*, *11*(3), 357–369.

Wolfram Research, I. (2016). Mathematica. Champaign, Illinois.