

*A1. Literature on which we draw*

Since the initial contribution of Lewis, evolutionary linguists and biologists studying models of language have developed a rich set of evolutionary models describing the distribution of languages, the emergence of words, syntax, and universal grammar, and the evolution of grammatical conventions such as verb regularization (Pagel (2009); Pagel et al. (2013); Komarova, Niyogi and Nowak (2001); Nowak, Plotkin and Jansen (2000); Lieberman et al. (2007); Ahern et al. (2016)). These models have a population of individuals who play languages, modelled as mappings from objects to signals, and who must successfully coordinate on a language in order to communicate. Trapa and Nowak (2000) show that evolutionarily stable languages are strict Nash equilibria where the number of signals are equal to the number of symbols, and each symbol is emitted with probability 1 conditional on the state. Empirically, studies of language evolution have documented “ultraconserved” words from the last ice age that remain in related forms in the current language distribution (Pagel et al. (2013)), and that much linguistic evolution happens in sharp, punctuated bursts (Atkinson et al. (2008)).

Plotkin and Nowak (2000) use results from coding theory to show that evolutionarily fit languages will efficiently transmit information. Efficiency is measured as the number of bits needed to transmit a message, and so Shannon’s coding theorem (Shannon (2001)) gives the theoretical upper bound on the efficiency of a channel, given the noisiness of the source encoding and decoding. In natural language, the noise is due to biological and cognitive constraints on signal processing. Christiansen, Chater and Culicover (2016) argues that even complex features of language, such as recursion, can be explained by adaptation to pre-given human constraints, and so the remarkable adaptiveness of humans to language is a result of natural selection of languages according to ease of acquisition and usefulness. Indeed, a recent study by Pellegrino, Coupé and Marsico (2011) finds that the information transmitted per second across spoken languages is quite stable, despite differences in speed of speaking and information per syllable. However, one thing lacking in these models is a justification for the persistence of “vague” or “ambiguous” linguistic conventions with significantly higher entropy than the optimal.

A second strand of literature looks at the economics and political economy of language and language policy. Within economics, the seminal paper is Lazear (1999) who makes the simple point that languages that increase the space of trading opportunities will be adopted, and makes numerous predictions that follow from this. The closest model to ours in Clingingsmith (2015), who models languages as conventions on networks. Clingingsmith shows that language growth follows a Gibrat’s law, and that the world’s languages are doubly-Pareto distributed. Also related is Laitin (2007), in that he models the choice of languages as a battle of the sexes, rather than pure coordination, type game.

*A2. Data on Linguistic Corpora*

We use the Google N-Grams corpus due to ease of access. However it is important to recognize severe limitations of this dataset, see Davies (2015) and criticisms by Pechenick, Danforth and Dodds (2015). An alternate source is the Corpus of Historical American English (COHA)<sup>1</sup>, “he or she” has increased in frequency as a share of 3-word phrases from less than 2 per million prior to 1970 to 15 per million in 1990.

<sup>1</sup>See Davies (2012). N-grams and word frequency data from the Corpus of Historical American English (COHA). Downloaded from <http://www.ngrams.info> on December 18, 2016. See <http://textlab.econ.columbia.edu/coha> for an interactive browser.

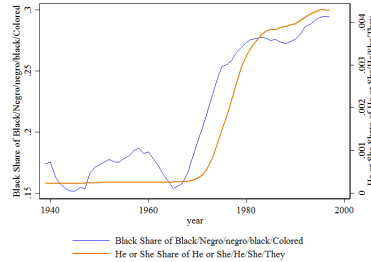


Figure A1. : Reproduced from text: Transitions in group titles: Relative increase in fraction Black and 3-word phrases that are “he or she” following 1960s.

### A3. Race and Gender Language Conventions

The terms for referring to race and gender identities have been transformed since the 1960s, largely due to attempts to shape the language made by activists and intellectuals. The transition from “colored” to “negro” to capitalized “Negro/Afro-American” to “Black/African-American” was not the result of spontaneous linguistic innovation, but instead deliberate campaigning. Booker T. Washington and the NAACP advocated for linguistic change in the early 20th century, with a victory scored when the New York Times imposed capitalized Negro in its style guides on March 7, 1930. Black power activists in the 1960s debated the merits of various terms and *Ebony* magazine joined them in promoting “Black”. This change was soon reflected in the overall culture, as can be seen from Figure A1.

Lerone Bennett, writing in *Ebony* magazine<sup>2</sup>, provides a historical perspective on racial terms in American English. “The movement for adoption of the word “Negro” was also given a strong impetus by militant radicals like W. E. B. Du Bois, who was one of the founders of the American Negro Academy, and militant nationalists like Marcus Garvey, who used the word “Negro” consistently and named his organization the Universal Negro Improvement Association.”

Gendered language has also been transformed, where the universal “he” has been curtailed in recent years. As with the transformation of racial language, the change was deliberately promoted by many independent deviations from the status quo (Pauwels (2003)). The magazine *Ms* titled to promote a new salutation of women was founded by two feminist activists Gloria Steinem and Dorothy Pittman Hughes. The year it was first published, 1972, the *Modern English Handbook* confirmed the traditional pronoun use: “He, alone, is usually preferred”, while modern guides often suggest “he or she” and suggest that gender not be presumed. A result of this norm is that gender-inclusive subjects have increased (as in Figure A1). (Curzan, 2014)

Deliberate attempts to change conventions back to inegalitarian ones sometimes fail. Mussolini’s fascist movement in Italy attempted to move Italians away from the pronoun *Lei*, which is honorific and to some fascists seemed effeminate, *lei* (with the lower case el) meaning “she”. The Partito Nazionale Fascista restricted the use of “Lei” among members pushing “voi” instead and then mandated the same practice for public employees in 1939. As can be seen from the figure A2, this reform was unsuccessful, with the secular relative decline in the use of “voi” only temporarily reversed during the two decades of Mussolini’s regime.

### A4. The model described informally in the paper

To describe a language, we let a message sending matrix  $P$  be a  $(|R|+|S|) \times |W|$  stochastic matrix mapping objects into words and a message receiving matrix  $P'$  be a  $|W| \times (|R| + |S|)$  stochastic matrix mapping words into object. Here,  $P_{ij}$  means that word  $j$  is emitted with probability  $P_{ij}$

<sup>2</sup>[http://www.virginia.edu/woodson/courses/aas10220\(spring2001\)/articles/names/bennett.htm](http://www.virginia.edu/woodson/courses/aas10220(spring2001)/articles/names/bennett.htm)

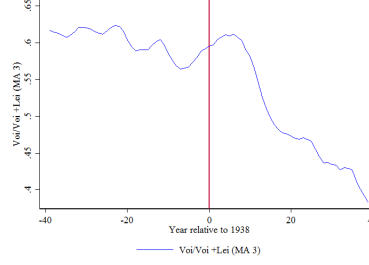


Figure A2. : Attempted deliberate reversal: Italian Voi/(Voi+Lei) around 1938.

when object  $i$  is being communicated and  $P'_{ji}$  means that object  $i$  is received when word  $j$  is heard. Thus  $P_{ij}P'_{ji}$  is the probability that two persons successfully communicate object  $i$  with word  $j$ . Then a language  $L$  is defined to be  $L := (P, P')$ . Thus the set of all languages,  $\mathcal{L}$ , is given by

$$\mathcal{L} := \Delta^{(|R|+|S|) \times |W|} \times \Delta^{|W| \times (|R|+|S|)}$$

To simplify, we consider  $L$  such that  $L = (P, P^T)$ , where  $P^T$  is the transpose of  $P$ . We then can regard a matrix  $P$  as a language. Under this assumption, the communication probability of two players using  $P$  and  $Q$  languages for object  $i$  using word  $j$  is given by  $P_{ij}Q_{ji}^T = P_{ij}Q_{ij}$ .

The payoff of an agent playing language  $P$  when communicating with an agent playing language  $Q$  is:

(A1)

$$U^A(P, Q) = \underbrace{\frac{1}{2} \sum_{i \in R} \sum_{j \in W} P_{ij} Q_{ij}}_{\text{Regular Interaction Sender}} + \underbrace{\frac{1}{2} \sum_{i \in R} \sum_{j \in W} P_{ij} Q_{ij}}_{\text{Regular Interaction Receiver}} + \underbrace{\frac{\theta \rho}{2} \sum_{i \in S} \sum_{j \in W} P_{ij} Q_{ij}}_{\text{Status Interaction Sender}} + \underbrace{\frac{\theta \rho}{2} \sum_{i \in S} \sum_{j \in W} P_{ij} Q_{ij}}_{\text{Status Interaction Receiver}}$$

(A2)

$$U^B(P, Q) = \underbrace{\frac{1}{2} \sum_{i \in R} \sum_{j \in W} P_{ij} Q_{ij}}_{\text{Regular Interaction Sender}} + \underbrace{\frac{1}{2} \sum_{i \in R} \sum_{j \in W} P_{ij} Q_{ij}}_{\text{Regular Interaction Receiver}} + \underbrace{\frac{(1-\theta)\rho}{2} \sum_{i \in S} \sum_{j \in W} P_{ij} Q_{ij}}_{\text{Status Interaction Sender}} + \underbrace{\frac{(1-\theta)\rho}{2} \sum_{i \in S} \sum_{j \in W} P_{ij} Q_{ij}}_{\text{Status Interaction Receiver}}$$

In the 2-state 1 symbol example in the paper,

$$P^u = \begin{pmatrix} & \text{V distinction} \\ \text{Plural} & 1-x \\ \text{Power} & x \end{pmatrix}, \quad P^e = \begin{pmatrix} & \text{V distinction} \\ \text{Plural} & 1 \\ \text{Power} & 0 \end{pmatrix}$$

and the resulting payoffs for Group A are:

$$(A3) \quad U^A(P^u, P^u) = (1-x)^2 + \rho \theta x^2$$

$$(A4) \quad U^A(P^e, P^u) = (1-x)$$

$$(A5) \quad U^A(P^u, P^e) = (1-x)$$

$$(A6) \quad U^A(P^e, P^e) = 1$$

and the payoffs for Group B are

$$(A7) \quad U^B(P^u, P^u) = (1-x)^2 + (1-\theta)\rho x^2$$

$$(A8) \quad U^B(P^e, P^u) = (1-x)$$

$$(A9) \quad U^B(P^u, P^e) = (1-x)$$

$$(A10) \quad U^B(P^e, P^e) = 1$$

We are interested in characterizing the stochastic stability of a language convention. Using methods we discuss in [Hwang, Naidu and Bowles \(2016\)](#) and in the Appendix, we find the resistances for each convention, which measure the relative difficulty of idiosyncratic behavior tipping the system from one equilibrium language to another. The resistance from  $P^u$  to  $P^e$ ,  $c(P^u, P^e)$ , and the resistance from  $P^e$  to  $P^u$ ,  $c(P^e, P^u)$  are given by

$$c(P^u, P^e) = \left[ \eta N^R \frac{(1+\rho(1-\theta))x-1}{(1+\rho(1-\theta))x} \right] \wedge \iota \left[ N^R \frac{(1+\rho\theta)x-1}{(1+\rho\theta)x} \right]$$

$$c(P^e, P^u) = \left[ N^R \frac{1}{(1+\rho\theta)x} \right] \wedge \iota \left[ \eta N^R \frac{1}{(1+\rho(1-\theta))x} \right].$$

From this we obtain the following theorem.

**THEOREM 1:** *Suppose that  $x > \frac{1}{1+\rho\theta}$ . We have the following characterizations:*

(i) (**Unintentional**  $\iota = 1$  **and equal population size**  $\eta = 1$ )  $P^u$  is stochastically stable if and only if  $\left[ N^R \frac{(1+\rho\theta)x-1}{(1+\rho\theta)x} \right] > \left[ N^R \frac{1}{(1+\rho(1-\theta))x} \right]$

(ii) (**Intentional**  $\iota = \infty$ )  $P^u$  is stochastically stable if and only if  $\left[ \eta N^R \frac{(1+\rho(1-\theta))x-1}{(1+\rho(1-\theta))x} \right] > \left[ N^R \frac{1}{(1+\rho\theta)x} \right]$

To show the first result in the text, we use Theorem 1 (i). From this, we obtain that convention  $P^u$  is stochastically stable if and only if

$$(A11) \quad x > \frac{1}{1+\rho\theta} + \frac{1}{1+\rho(1-\theta)}$$

To study the effect of  $\theta, \rho$  on the stochastic stability of convention  $P^u$ , we first observe that the right hand side of the inequality in (A11) is decreasing in  $\rho$  and  $\theta$ . Also, obviously a higher value of  $x$  is more likely to satisfy the inequality in (A11). Thus, the higher  $\theta, \rho$ , and  $x$ , the more stochastically stable  $P^u$  convention. The second result in the text follows from (ii) of Theorem 1. That is, the larger  $\eta$ , the more likely  $P^u$  (the convention favored by the group  $B$ ) is stochastically stable.

#### REFERENCES

\*

#### REFERENCES

- Ahern, Christopher A, Mitchell G Newberry, Robin Clark, and Joshua B Plotkin. 2016. "Evolutionary forces in language change." *arXiv preprint arXiv:1608.00938*.
- Atkinson, Quentin D, Andrew Meade, Chris Venditti, Simon J Greenhill, and Mark Pagel. 2008. "Languages evolve in punctuational bursts." *Science*, 319(5863): 588–588.

- Christiansen, Morten H, Nick Chater, and Peter W Culicover.** 2016. *Creating language: Integrating evolution, acquisition, and processing*. MIT Press.
- Clingingsmith, David.** 2015. “Are The World’s Languages Consolidating? The Dynamics and Distribution of Language Populations.” *The Economic Journal*.
- Curzan, Anne.** 2014. *Fixing English: Prescriptivism and language history*. Cambridge University Press.
- Davies, Mark.** 2012. “The 400 million word corpus of historical American English (1810–2009).” Vol. 325, 231.
- Davies, Mark.** 2015. “The importance of robust corpora in providing more realistic descriptions of variation in English grammar.” *Linguistics Vanguard*, 1(1): 305–312.
- Hwang, Sung-Ha, Suresh Naidu, and Samuel Bowles.** 2016. “Social conflict and the evolution of unequal conventions.” Santa Fe Institute Working Paper.
- Komarova, Natalia L, Partha Niyogi, and Martin A Nowak.** 2001. “The evolutionary dynamics of grammar acquisition.” *Journal of theoretical biology*, 209(1): 43–59.
- Laitin, David D.** 2007. *Language repertoires and state construction in Africa*. Cambridge University Press.
- Lazear, Edward P.** 1999. “Culture and Language.” *Journal of Political Economy*, 107(S6): S95–S126.
- Lieberman, Erez, Jean-Baptiste Michel, Joe Jackson, Tina Tang, and Martin A Nowak.** 2007. “Quantifying the evolutionary dynamics of language.” *Nature*, 449(7163): 713–716.
- Nowak, Martin A, Joshua B Plotkin, and Vincent AA Jansen.** 2000. “The evolution of syntactic communication.” *Nature*, 404(6777): 495–498.
- Pagel, Mark.** 2009. “Human language as a culturally transmitted replicator.” *Nature Reviews Genetics*, 10(6): 405–415.
- Pagel, Mark, Quentin D Atkinson, Andreea S Calude, and Andrew Meade.** 2013. “Ultraconserved words point to deep language ancestry across Eurasia.” *Proceedings of the National Academy of Sciences*, 110(21): 8471–8476.
- Pauwels, Anne.** 2003. “24 Linguistic Sexism and Feminist Linguistic Activism.” *The handbook of language and gender*, 550.
- Pechenick, Eitan Adam, Christopher M Danforth, and Peter Sheridan Dodds.** 2015. “Characterizing the Google Books corpus: Strong limits to inferences of socio-cultural and linguistic evolution.” *PloS one*, 10(10): e0137041.
- Pellegrino, François, Christophe Coupé, and Egidio Marsico.** 2011. “Across-language perspective on speech information rate.” *Language*, 87(3): 539–558.
- Plotkin, Joshua B, and Martin A Nowak.** 2000. “Language evolution and information theory.” *Journal of Theoretical Biology*, 205(1): 147–159.
- Shannon, Claude Elwood.** 2001. “A mathematical theory of communication (reprinted with corrections from 1948 version).” *ACM SIGMOBILE Mobile Computing and Communications Review*, 5(1): 3–55.
- Trapa, Peter E, and Martin A Nowak.** 2000. “Nash equilibria for an evolutionary language game.” *Journal of mathematical biology*, 41(2): 172–188.