The Evolution of Egalitarian Socio-Linguistic Conventions

By Suresh Naidu, Sung-Ha Hwang, and Samuel Bowles[*]

Half a century ago two works laid the foundation for modern sociolinguistics. Lewis (1960) studied a speech community using a system of symbols and grammatical rules and coordinating on the mapping from symbols to states of the world in order to communicate effectively. Economists and others built on Lewis' work to erect the modern theory of conventions, their persistence and occasional transformations (Young (1998)). Subsequent research has shown that languages that are likely to emerge and persist in an evolutionary dynamic are informationally efficient under the constraints imposed by human cognitive and sensory systems.

At about the same time Brown and Gilman published their "Pronouns of Power and Solidarity" exploring the fact that "a man's consistent pronoun style gives away his class position" and that this "power semantic" had been the norm in many Indo-European languages for at least half a millennium despite recurrent contestation by egalitarian language innovators (Brown and Gilman (1960))

We focus in this paper on what Brown and Gilman called the T-V distinctions (e.g. the familiar "Tu" vs the formal "Vous" in French), the semantics of which typically involve an ambiguity in that the V pronoun may denote superordinate status as well as plurality (Tabellini (2008)). The T-V status markers are far from unique as linguistic features of group interests and identity. Labov (2011) and other sociolinguists have established, for example, that class-based accents are pervasive. Linguistic markers of unequal status with respect to race and gender are also common (as we show in our online Appendix.)

Here we use recently developed models of conventions to model the evolution of the non reciprocal power semantic studied by Brown and Gilman. Our objective is to provide a framework that is consistent both with the long term persistence of the T-V distinction (despite the ambiguity intrinsic to the dual use of the V pronoun for status and plurality) and with its recent displacement in many languages by an egalitarian pronoun convention. To do this we consider a population composed of two classes in which communication of relative status imposes subjective costs on the subordinate class and also may serve as a socially valuable coordination device.

As a benchmark, we specify conditions in this battle of the sexes coordination game environment under which evolutionarily successful languages are unambiguous and egalitarian, eschewing pronominal markers of status. But we also show that if, as is often the case, the subordinate population is large relative to the elite and linguistic innovations are intentional (rather than mutation-like accidents) then ambiguous and unequal conventions are likely to emerge and persist over a long period, consistent with the history of the T-V conventions.

I. Intentional Linguistic Changes

We use changes in T-V distinctions of status when used as a singular pronoun as an example of bottom-up intentional linguistic changes towards egalitarian linguistic conventions. The honorific (non-reciprocal) usage where superiors are addressed by "V" and subordinates by "T" has been in decline. According to Kachru and Smith (2008) "there is an increasing tendency to address all intimates, regardless of status, with the T-pronoun, and all strangers with the V-pronoun," a trend identified Brown and Gilman in their initial paper. We focus on political transitions, where changes in payoffs to the communication of status may have made transitions between linguistic conventions more likely.

A good example is French, as shown in Figure I, which plots the relative frequency of

"Vous" to "Tu" (case-insensitive) in the Google N-grams database. While this is naturally noisy and imperfect data[1], it shows a relative increase in the "Tu" form around the French Revolution. While we cannot distinguish between formal and informal symmetric (vs asymmetric use), this is consistent with revolutionary norms of egalitarianism that prevailed during the Revolution.

The revolutionary Committee for the Public Safety denounced "vous" as a feudal anachronism; Robespierre *tu*'d the Assembly's President. Anderson (2007) writes "the idea of using *tu* in all circumstances was first proposed in an article in the *Mercure National* on December 14, 1790....No laws were passed registering the mandatory use of *tu* but...[it]..began to spread. Now the baker's apprentice could address his master and clients in a familiar form, a practice that had been strictly forbidden." The N-gram data in the figure also show a sharp change following the student movements of 1968, with activists again deliberately using "Tu" to address superiors.
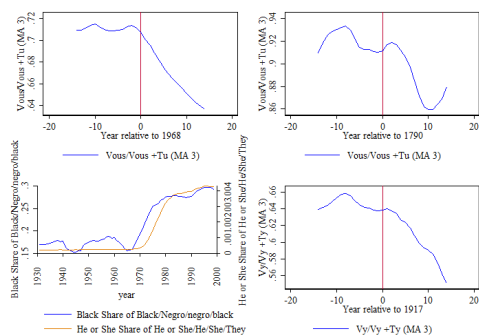

Figure 1. : Clockwise from top left 15 year windows around political events: Vous/(Tu+Vous) relative frequency in the 2012 French Google N-grams corpus around 1968 and 1790. Russian Vy/(Vy + Ty) around 1917. Transitions in group titles: Relative increase in fraction Black and 3-word phrases that are "he or she" following 1960s.

The Russian revolution provides another example as can be seen in the Russian corpus in Figure I. A demand of revolutionary workers during the Lena strike of 1912 was to be addressed in the polite mode Stites (1988). The use of the honorific within the Soviet army was abolished in 1917. During and after the revolution Russian intellectuals and activists intentionally began using the informal mode of address universally. These challenges to the convention diffused into a larger linguistic change (Corbett, 1976).

Deliberate bottom up challenges to status quo conventions sometimes fail. The 17th century Society of Friends (Quakers) raised the banner of Plain Speech, according to which the informal "Thou" or "Thee" was prescribed for all social interactions rather than the asymmetric formal "You" or reverential "Ye" (innovations introduced by nobility after the Norman invasion). Fox wrote: "... when the the Lord sent me forth into the world, He forbade me to put off my hat to any, high or low: and I was required to Thee and Thou all men and women without any respect to rich or poor, great or small." But little came of it, and the informal pronouns eventually were all but abandoned, while "you" lost its status connotations. But these egalitarian challenges also sometimes succeed, as shown in the changing race and gender terms in Figure I, which we discuss more fully in the Appendix.

II.   An Evolutionary Linguistic Model

Our model is a contribution to an evolutionary socio-linguistics that draws on both our previous work on intentional evolutionary equilibrium selection and a rich literature in evolutionary linguistics, which we cite more fully in the Appendix. Like Lewis, we represent a language as a convention, that is, a mutual best response of speakers who may adopt differing languages. By the evolutionary success of a language convention we mean roughly the likelihood over a very long period of time that a population will coordinate on that particular convention when speakers typically best respond by conforming to the status quo convention but occasionally (with probability ϵ) innovate, responding idiosyncratically and deviating from the convention.[2]

---

[1] The frequencies are calculated from the universe of books digitized by Google, see Michel (2011) and discussion in Appendix.

[2] More technically an evolutionarily successful language is stochastically stable. A Nash equilibrium (for example a language convention) is *stochastically stable* if the resulting ergodic distribution of strategies in the population puts positive

We model language evolution as decentralized process in which the common-language coordination outcome may occur as an unintended emergent property of uncoordinated interactions, not a result mandated by some central authority. But while linguistic outcomes are unintended, language behaviors – conformity or resistance to a convention – are deliberate. When individuals deviate from the prevailing convention, they do so, not in error, but intentionally, adopting an alternative convention in which the would be better of, were the rest of their class to do the same.

We build on the model of language of Nowak, Plotkin and Krakauer (1999), extending it in a number of ways. We incorporate two populations, A and B, where both population sizes N_A and N_B are large and $N_A = \eta N_B$, where η is the relative group size of A. We assume that the payoffs are asymmetric so as to capture a battle of the sexes logic: a common language convention is preferred by both populations, but they differ on which convention they prefer.

Members of the A population are randomly paired with B's and may with equal probability be a sender or a receiver. A language strategy is a probability matrix mapping objects to symbols (the sending matrix), and the transpose of that matrix is the "receiver" matrix that decodes symbols back into objects. For example, a sender who utters "letter" could with some probability intend "one of the items making up the alphabet" and with the complementary probability "a written message."

Communication is successful if the object signalled by the sender is decoded correctly by the receiver. Since the receiver matrix is the transpose of the sending matrix, communication occurs with highest probability when both agents are coordinating on the same language. Unsuccessful communication gets a payoff of 0.

We divide the space of objects to be communicated into "Regular" (R) and "Status-relevant" (S). Agents get a payoff of 1 from successfully communicating the R objects. When communicating about the Status relevant objects there is some total payoff to successfully communicating status differences. For example passage through a doorway may be coordinated by the norm that the higher status person goes first,

and the benefits of observing this norm (avoiding collisions or endless deferring to the other) may be communicated by some aspect of the language on which the two coordinate, such as the T-V distinction.

We denote this total benefit by ρ, and we imagine systems of economic and social interaction in which this might be a considerable magnitude. For example, if costly conflicts are sometimes avoided by the mutual recognition of status differences, as with the rule that a subordinate must cede to a dominant (observed in many primates), then ρ would be substantial.

The member of group A gets payoff $\theta\rho < 1$, while the B member of the interaction gets payoffs $(1-\theta)\rho$, where $\theta < \frac{1}{2}$. Thus group A is the relatively "low status" group. In the above doorway example the A member derives some benefit by avoiding delay or a collision, but at the cost of publicly acknowledging a socially inferior status.

We consider just two languages in which we let P^e be an egalitarian language that does not let an agent communicate status and P^u be an inegalitarian language that does let an agent communicate status with with probability x. So if the sender using P^u approaching the doorway utters "I will see you (using a V pronoun) later," it could mean with probability $1 - x$ the plural "you and your family" (thereby not conveying status and thus forgoing coordinating on efficient passage through the door) or the singular "you, my recognized superior" with probability x (successfully avoiding delay or collision in the doorway).

Thus, because in P^u V may designate either plurality or status, there is a probability of miscommunication even when both players are using P^u. In this case both agents using P^u will understand each other with probability $(1 - x)^2$ when communicating the Regular object and with probability x^2 when communicating the Status object. We assume 2 strict Nash equilibria, implying payoff restrictions that make x a monotone increasing measure of clarity (lack of ambiguity) in communicating status (i.e. $\frac{1}{2} < \frac{1}{1+\rho\theta} < x < 1$).

By contrast, both agents using P^e will never mis-communicate plurality, but but at a price: they will also never communicate status even when that is relevant. When an agent using

mass on that equilibrium as $\epsilon \to 0$.

P^e encounters an agent using P^u they receive only the payoff from communicating plurality, and then only when the agent using P^u intends to communicate plurality. Both agents using P^e will always communicate the Regular object correctly, obtaining payoffs 1. When an agent using P^e encounters an agent playing P^u (or agent using P^u encounters an agent using P^u), they only receive payoff 1 from communicating plurality, which occurs with probability $1 - x$.

Finally, both agents using P^u will understand each other only with probability $(1 - x)^2$ when communicating the Regular object and x^2 when communicating the Status object. Thus both agents obtain payoff 1 from communicating the Regular object with probability $(1 - x)^2$, while group A agents obtain payoff $\rho\theta$ and group B agents obtain payoff $\rho(1 - \theta)$ each with probability x^2. In the Appendix, we derive the payoffs from language coordination and miscoordination more formally, and show that we can represent the payoffs as a simple coordination game in Table 1, with $U_u^A = (1-x)^2 + \rho\theta x^2$ and $U_u^B = (1-x)^2 + \rho(1-\theta)x^2$.

Group B

		P^e	P^u
Group A	P^e	$1, 1$	$1 - x, 1 - x$
	P^u	$1 - x, 1 - x$	U_u^A, U_u^B

Table 1—: **Payoffs in the Language Game.**

III. Persistence and Change

We now turn to determining which equilibrium is selected under an explicit evolutionary dynamic. We impose the dynamics in Hwang, Naidu and Bowles (2016)[3], where agents myopically play best responses to the previous period's distribution of strategies in the population, and have the opportunity to play idiosyncratically, deviating from the convention, with probability ϵ.

Transitions from from one convention to another occur when the number of deviants from the status quo convention in one class is sufficient to induce the best responding members of the other class to adopt the alternative convention. This dynamical process is ergodic, so that

[3]When $x = 1$ this game reduces to the game studied in that paper.

the population never gets 'locked into' one language or the other, but will use one convention for a long time before making a transition to the other, only to return after another long period of stasis.

To capture the purposeful nature of language change we say that when agents have the opportunity to play idiosyncratically they are less likely to deviate from the convention when the status quo convention is one they prefer. The parameter ι measures the degree to which innovations are "intentional": supported only on the strategies that if widely adopted by their class would result in a transition to an equilibrium that is better for them than the current one. A concrete interpretation of ι is that intentional innovation is greater where those who are disadvantaged by the current non-reciprocal power semantic are aware of their collective inferiority and are willing even at a cost to themselves to challenge it by deviating from its asymmetric pronoun use.

Besides varying the degree of intentionality ι, we allow for unequal population sizes parameterized by η above. As in Hwang, Naidu and Bowles (2016), these modifications allow stochastic evolutionary game theory to be used to model equilibrium selection in environments where there are conflicts of interests between groups that differ in size, payoffs, and level of internal organization and mobilization.

Some of our results are unsurprising. Across parameter values, the unequal language will be relatively more persistent, the more valuable is the communication of status differences (the greater is ρ) and the lesser the ambiguity in accomplishing this (the greater is x). We also find that when innovation is intentional a population that has a higher rate of idiosyncratic play (equivalent to lower η) will be favored and will spend more time speaking the convention they prefer. This unsurprising result appears consistent with recent changes in gender and racial language especially in light of the fact (see the appendix) that in these cases media and government played important roles in amplifying the influence of a few innovators.

Other results are more counter intuitive. The dynamic we have modeled, like that in Young (1998) favors languages that are egalitarian: a unequal language (that is with a small θ) with a higher ρ may be less persistent than a more

equal variant of P^u with higher θ.

The intuition behind this result (shown formally in the appendix) is that it takes more idiosyncratic play to dislodge a language convention the more egalitarian it is. Suppose a population is using the language without status distinctions and that were a transition to an unequal convention to occur, the subordinate group in the population would receive virtually no benefits from the communication of status (low θ). Then it would take almost all of the privileged group to play idiosyncratically to make adopting the unequal language a best response for the subordinate group. Because this would be an extraordinarily rare event, the more unequal is P^u, the more time will the population spend at the P^e convention.

Yet this case may not pertain to historically relevant environments. When linguistic innovations are directed (ι large) and there is a relatively large subordinate class (η large), idiosyncratic play can stabilize a convention that is both unequal (θ small) and maximally ambiguous (x close to $\frac{1}{2}$) even where communicating status differences is of little value (ρ small).

The intuition behind this is that transitions are induced by extreme realizations of the stochastic process generating innovations in which a large fraction of a given population does not best respond. These extreme realizations are more likely in small populations for the same reason that the variance around a sample mean is greater, the smaller is the sample.

This makes it more likely that it is the members of a small population whose innovations will induce a transition. But if their innovations are random rather than intentional, they will as likely induce a transition away from their preferred convention as towards it.

However, if people innovate intentionally their idiosyncratic play induces transitions only in a direction from which they benefit. Under these conditions a small group is advantaged: when deliberately innovating it can easily destabilize an unfavorable status quo convention.

Thus, evolutionarily successful linguistic conventions – including the non reciprocal power semantic studied here – need not excel in clarity of communication, as the long persistence of the asymmetrical use of T-V pronouns suggests.[4]

---

[4]In Hwang, Naidu and Bowles (2016) we show how these

REFERENCES

Anderson, James M. 2007. *Daily life during the French Revolution.* Greenwood Publishing Group.

Brown, Roger, and Albert Gilman. 1960. "The pronouns of power and solidarity. Style in language, ed. by Thomas A. Sebeok, 253-76."

Corbett, GC. 1976. "Address in Russian." *Journal of Russian Studies*, , (31): 3–15.

Hwang, Sung-Ha, Suresh Naidu, and Samuel Bowles. 2016. "Social conflict and the evolution of unequal conventions." Santa Fe Institute Working Paper.

Kachru, Yamuna, and Larry E Smith. 2008. *Cultures, contexts, and world Englishes.* Routledge.

Labov, William. 2011. *Principles of linguistic change, cognitive and cultural factors.* Vol. 3, John Wiley & Sons.

Lewis, David. 1960. *Convention: A philosophical study.* John Wiley & Sons.

Michel, Jean-Baptiste et al. 2011. "Quantitative analysis of culture using millions of digitized books." *science*, 331(6014): 176–182.

Nowak, Martin A, Joshua B Plotkin, and David C Krakauer. 1999. "The evolutionary language game." *Journal of Theoretical Biology*, 200(2): 147–162.

Stites, Richard. 1988. *Revolutionary dreams: Utopian vision and experimental life in the Russian revolution.* Oxford University Press.

Tabellini, Guido. 2008. "Institutions and Culture." *Journal of the European Economic Association*, 6(2-3): 255–294.

Young, H. Peyton. 1998. *Individual Strategy and Social Structure: An Evolutionary Theory of Institutions.* Princeton University Press.

results extend to games played on arbitrary networks. We provide conditions on the topology of arbitrary networks that can make P^u the more persistent, whereby a small well-connected set of A's can induce a large set of B's to change their language with only a few idiosyncratic innovations.