Linguistic exploitation

JONATHAN POOL

"Not the gun but the word is the symbol of authority. The most frequent governmental activities are talking, writing, listening, and reading" (Lindblom 1977: 52). If Lindblom is right about what governments do, then the word is not merely a *symbol* but also a *tool* of authority. If words are tools of authority, then language, which allows people to produce words, must be a tool for making tools of authority. And, if language is a machine tool in the authority industry, then we should also expect people to compete for control over language.

As expected, we find struggles over language taking place every day. Political activists devote much of their energy to such struggles. And the nature of politics may be influenced by the fact that linguistic competition is one of the determinants of political success. According to one interpretation, suggested by Edelman (1977), those who have political power use it to get power over language, and those who have power over language use it to get political power, with the result that the ideal of democratic government is never achieved. Myers-Scotton (1989) has offered a variant of this argument, under the heading of "elite closure," which she defines as

> a tactic of boundary maintenance: it involves institutionalizing the linguistic patterns of the elite, either through official policy or informally-established usage norms. This limits access to socioeconomic mobility and political power to those societal members who possess the requisite linguistic patterns of the elite.

One tactic in struggles over language is to obtain control over the definition of *correctness*. Correctness in language, as in anything, is scarce. It is an attribute of only some languages and some linguistic forms. If all languages and forms had it, we would not even have a concept of correctness.

It would be naïve to suppose that the distinctions between what is linguistically correct and what is linguistically incorrect arise without

0165–2516/93/0103–0031 $2.00

Int'l. J. Soc. Lang. 103 (1993), pp. 31–55

deliberate manipulation. Many people have an interest in influencing these distinctions. Defining certain forms of languages as correct can serve the interests of certain persons, because they know those forms or languages, or are identified with them, or are liked by their speakers, or for various other reasons. Norms of correctness often determine how respectfully the speakers of a language address or refer to members of certain groups, such as males, females, and the elderly.

Most people seem to accept norms of correctness as natural, regardless of who has designed them, for what purpose, and with what effects on the distribution of power. One reason for acceptance may be a disinclination to believe that the origin of these norms is deliberate choice. It is often presumed that norms of correctness arise through unconscious, uncoordinated actions of many people as a matter of convention. In an attempt to explain the moral coloring to the politics of language, it has been hypothesized that language is naturally or primievally an evaluative domain. Haugen (1966: 288–289) asserts that language cannot be left to individual discretion: people need to be told what is right and wrong in language.

Even if people thought that certain powerful persons had decided what forms and languages are correct, people would not necessarily object to these decisions. Myers-Scotton (1989) finds that positive popular attitudes toward elite language forms is a "well-documented sociolinguistic universal." Perhaps this deference is entirely voluntary; according to Jespersen (1964 [1925]: 85),

> it is characteristic of human nature that most people wish for an external authority, even in linguistic questions. Just as they uncritically adapt themselves to much of what their tailors or their newspapers tell them about the particular cut of clothes which "people" are wearing at the moment, so they wish to have some definite direction as to the pronunciation, spelling and use of words. If they have not a teacher at hand who can give them an infallible rule to settle their doubts, they rely blindly on the dictionary or grammar which they happen to have.

Jespersen's account implies that people are more comfortable obeying than innovating. His explanation is that people simply have a taste for compliance.

The other face of compliance, however, is enforcement. Perhaps people also have a taste for forcing others to comply with standards of correctness, and for punishing those who do not comply. Durkheim (1964 [1893]: chap. 2) argues that societies define rules of behavior beyond those that serve the societies' interests. Once they are defined, he says these rules must be enforced, not to inflict gratuitous harm, and not even to improve

or deter deviant behavior, but rather to "maintain social cohesion," which is damaged by any unpunished violation of society's laws even though the act would have been considered innocuous if the law prohibiting it had not existed (Durkheim 1964 [1893]: 106–108). Durkheim's hypothesis could be stretched to say that a society can maintain its solidarity only by having deviants with whom its compliant core members can contrast themselves. To guarantee a supply of deviants, the society may need to have arbitrarily difficult rules.

Language may offer a convenient arena for the definition and enforcement of arbitrarily difficult societal rules. In the normal use of language, any incorrectness is witnessed by persons other than the perpetrator, and the perpetrator's identity is not in doubt. Compliance with linguistic rules is, then, easily monitorable. The complexity of language also makes it easy to define rules that only some of the population will be able to obey. The little industry of books and magazine columns devoted to linguistic criticism and advice (such as works by Mario Pei and Edwin Newman in the United States and Pierre Beaudry and Pierre Baudruche in Quebec) testifies to the difficulty of linguistic rules. The fact that millions of people violate rules of correctness after studying them for years in schools points to the same conclusion.

In the struggles to determine which words, rules, and languages are correct, we find two apparently opposite patterns. People seek to define a standard of correctness and *make* everyone adhere to it. Yet, people also seek to define a standard of correctness and *prevent* others from becoming competent in it. Definitions of correctness can serve either of these goals. If the correct form or language is simple and public, anyone can learn it, and everyone knows what is to be learned. If it is complex and arcane, it is costly to master, and those who have power over language preserve discretion in assessing whether a given utterance conforms to the standard. What is called a "standard" or "official" language can exhibit either of these tendencies or a mixture of them (cf. Das Gupta and Gumperz 1968: 154).

The apparently contradictory aims of linguistic standardization lead me to propose that we model the choices of those who have control over language. How difficult will they choose to make the achievement of linguistic correctness? What conditions will lead the authorities to make a linguistic standard accessible or exclusive?

One way to approach this problem is through a model of "predatory rule," following Levi (1983). Consider the following assumptions:

Assumption 1. A *polity* with a continuum of *citizens* is partitioned into two *classes*: the *ruling class* (*rulers*) and the *subject class* (*subjects*). The subject class produces an amount of *wealth* proportional to its size, and the ruling class *extracts* a share of this wealth. The extracted wealth is allocated uniformly to the ruling class. The wealth that is not extracted is allocated uniformly to the subject class.

Assumption 2. The *technology of rule* is the maximum proportion of the produced wealth that can be extracted from the subject class. The technology of rule is positive and less than 1.

Assumption 3. The *compliance level* is the proportion of the maximum extractable wealth that is actually extracted. The compliance level is 0 when the ruling class is smaller than its *minimal size*; is 1 when the ruling class is larger than its *effective size*; and otherwise varies with the size of the ruling class directly and linearly from 0 to 1. The minimal and effective sizes of the ruling class are positive and less than the size of the polity. The effective size is larger than the minimal size.

Assumption 4. There is a *language of rule*. The ruling class consists of all citizens who choose to *master* the language of rule.

Assumption 5. The language of rule has some positive *difficulty*. If the ruling class is not empty, it determines the difficulty of the language of rule by majority vote.

Assumption 6. All citizens have *linguistic resistance*, which impedes their mastery of the language of rule. Linguistic resistance is distributed uniformly in the polity between 0 and a positive maximum rate.

Assumption 7. Each ruler incurs *linguistic cost* at a rate equal to the difficulty of the language of rule, multiplied by the ruler's rate of linguistic resistance.

In the polity modeled here, economic and linguistic power are united in the ruling class. The economic power of the ruling class lies in its ability, if it is large enough, to extract some of the subjects' wealth. The linguistic power of the ruling class lies in its right to define the difficulty of the language of rule, which must be mastered by any citizen who joins the ruling class.

The ruling class's economic and linguistic powers are limited, however. There is a maximum extractable proportion of the subject class's wealth, called the technology of rule, and it is beyond the ruling class's control. The actual extracted wealth does not necessarily reach its maximum: it varies with the size of the ruling class. To regulate its size, the ruling class can use only one tool: defining the difficulty of the language of rule. And there is a constitutional limit on this linguistic power: the ruling class must be internally democratic, using majority voting.

The cost of mastering the language of rule is assumed to vary from citizen to citizen, because people differ in linguistic aptitude. Assumption 6 defines aptitude negatively, as resistance (lack of aptitude). To simplify the analysis, I have assumed that linguistic resistance has a minimum rate of 0 and is distributed uniformly between 0 and its maximum rate. For example, if the maximum rate is 10, then exactly 10 percent of the citizens have resistance rates between 0 and 1, another 10 percent have rates between 1 and 2, and so on (cf. Selten and Pool 1982).

In most real polities, each citizen constitutes only a small fraction of the population. If one citizen changed classes, the impact on the sizes of the classes would be too small to notice. My model carries this tendency to its extreme by assuming that the polity is a *continuum.* It contains an infinite number of citizens. This assumption implies that each citizen is only an infinitesimal fraction of the population, enjoys only an infinitesimal amount of wealth, exhibits only an infinitesimal amount of linguistic resistance, and incurs only an infinitesimal amount of linguistic cost. I have resorted to this idealization because it simplifies the deduction of conclusions. With a continuun of citizens, any subset of citizens constituting a positive fraction of the population can enjoy a positive amount of wealth, but a single citizen can only be said to receive wealth at a positive *rate.* The same holds for linguistic resistance and linguistic cost.

Even with this simplified picture of a two-class polity where the ruling class has complete power over the language of rule, the model's implications need to be examined with care. Would the ruling class ever make the language of rule so difficult that all rulers would decide to defect from their class, creating an anarchy? Could the language become trivially easy to master, so that there is almost no cost attached to doing so? Would all citizens become rulers, in that case? What happens to the language of rule when technological progress raises the technology of rule? Are there outcomes that, once reached, remain in effect indefinitely, or does every adjustment of the difficulty of the language of rule cause changes in the composition of the ruling class, leading to further revisions of the language when the next vote of the rulers is taken? If there are stable outcomes, how do they compare with the outcomes in a polity where the ruling class decides on its size directly, rather than by manipulating a language of rule?

Common sense may suggest answers to some of these questions. My common sense tells me that the language requirement should reduce the number of rulers in a polity. The linguistic barrier to entry into the ruling class makes membership in that class more costly than it would otherwise be. Consequently, I would expect some citizens who might have entered the ruling class without a language requirement to remain subjects because a language requirement is in force. With a reduced ruling class, I would expect the subject class to keep more of its wealth than it would if the ruling class could simply set its own size. Furthermore, the fact that it is the rulers rather than the subjects who must incur the cost of mastering a non-native language seems to be an obvious egalitarian element of the regime. Thus, the language of rule should serve as a limit on the exploitative impact of the ruling class. But are these intuitions supported by precise analysis?

For answers we need more specificity about how the choices are selected and what combination of choices will be accepted as the "outcome." I shall predict that the outcome will satisfy five principles:

Prediction principle 1. Sincere class choice.
Prediction principle 2. A directional agenda.
Prediction principle 3. Incrementally sincere voting.
Prediction principle 4. An equilibrium of class choices.
Prediction principle 5. An equilibrium of voting results.

The principle of *sincere class choice* says that any citizen chooses to join the class that maximizes that citizen's rate of *net wealth* under *current conditions.* A subject's rate of net wealth is simply the rate of wealth allocated to that subject. A ruler's rate of net wealth is the rate of wealth allocated to that ruler, reduced by the ruler's rate of linguistic cost. In determining which class maximizes one's rate of net wealth, one considers only current conditions, namely (1) the current difficulty of the language of rule and (2) the current class choices of all other citizens. Such a choice is called "sincere" because it directly expresses the citizen's assumed preference for a maximum rate of net wealth. What this principle excludes is a sophisticated class choice that might lead to a lower rate of net wealth under current conditions, motivated by the belief that the choice would help to change the conditions. For example, a citizen whose rate of net wealth is lower in the ruling class might nevertheless join that class in order to get the opportunity to vote for an easier language of rule, thereby contributing to conditions that would later raise that citizen's rate of net wealth. Such sophisticated class choices are excluded by this principle.

The principle of a *directional agenda* says that the ruling class votes only on proposals to change the difficulty of the language of rule in a particular direction (up or down). Particular amounts of the change are not voted on.

The principle of *incrementally sincere voting* says that each ruler always votes to change the difficulty of the language of rule in some direction if and only if every change in that direction smaller than some amount would *foreseeably* increase that ruler's rate of net wealth. Thus, all voting choices are based on the impact of minimal, or incremental, changes. If a slightly more difficult language of rule would reduce one's net wealth rate, but a much more difficult language would raise one's net wealth rate, this principle says that one votes against an increase in difficulty. Further, the only impacts on one's rate of net wealth that one considers are foreseeable impacts. These are the impacts that would follow from the exercise of sincere class choice by all citizens, without regard to the outcomes of any future votes. Thus, sincere voting requires rulers to consider how a slightly more difficult or less difficult language of rule would alter the size of the ruling class, how this change in size would alter the tax base and the compliance level, and how these changes, combined with the ruler's changed linguistic cost, would affect the ruler's own rate of net wealth. But the ruler is not permitted to look farther ahead to anticipate how the expanded or compacted ruling class would next vote.

The principle of an *equilibrium of class choices* says that the class choices of all citizens must be best replies to one another. In other words, any citizen who changes classes while all other citizens remain in their current classes must fail to obtain an increased rate of net wealth.

The principle of an *equilibrium of voting results* says that a proposed increase and a proposed decrease in the difficulty of the language of rule must both fail to get majority votes in the ruling class. I shall refer to outcomes satisfying *both* of these equilibrium principles as *stable* outcomes.

Like the seven original assumptions, the prediction principles are somewhat unrealistic, but the unrealism may be less serious than it appears to be. Sincere class choice and sincere voting require citizens to know all quantities and functions in the modeled polity, including one another's linguistic resistance rates and the compliance level as a function of the size of the ruling class. This would be a surprising amount of shared knowledge, but the equilibrium principle makes most of this knowledge unnecessary. At equilibrium, a citizen needs to know only which of two class choices is better, and a ruler additionally needs to know whether it is better to slightly raise, slightly lower, or leave unchanged the difficulty of the language of rule. Such knowledge need not result from calculations. Citizens might reach an equilibrium through trial and error, experiencing changing levels of net wealth as they change classes and as the rulers make small changes in the difficulty of the language of rule. Such experimentation would violate the principles of sincere class choice and sincere voting but might lead to the same outcomes. What kinds of trial-and-error processes can generate the predicted outcome is a question for further study.

The equilibrium principle could also be attacked as unrealistic, since it seems to imply that the linguistic costs incurred by rulers are adjusted whenever the difficulty of the language of rule changes. In reality, it would seem that rulers who learn a difficult language have paid a sunk cost and cannot recover any of it if they subsequently vote to make the language easier. Consequently, a language of rule might be kept more difficult than the equilibrium principle would predict. But the unrealism of this principle may be partly superficial. The cost of mastering a language is to some extent a recurring one. People forget what they know, so learning a language is a continuing burden. Users of difficult languages incur repeated costs, as many writers of English are reminded when they make frequent pauses to look up spellings in a dictionary. In addition, rulers usually want to pass their power to their descendants and may perceive their children's language-learning costs as their own.

The prediction principles, whatever their realism, make the model precise enough so that outcomes can be predicted. It will turn out that the model does not lead to the conclusion that I guessed on the basis of my common sense. Under some conditions the kind of regime modeled here, which I shall call a *linguistic regime*, is characterized by a *larger* ruling class than a regime where the rulers have the power to set their number merely by making a decision (I shall call this a *fiat regime*). Other results, which escaped my intuition entirely, also emerge. Here are some of the predictions the model makes:

Result 1. The subject class is not empty.
Result 2. The ruling class may be empty.
Result 3. No ruler has a greater rate of linguistic resistance than any subject.
Result 4. All subjects have the same rate of net wealth as the least wealthy ruler, if the ruling class is not empty.
Result 5. The ruling class is no larger than its effective size and no smaller than its minimal size, if it is not empty.
Result 6. The ruling class is a minority of the population.
Result 7. The size of the ruling class, if it is not empty, increases with the technology of rule in a linguistic regime but is independent of the technology of rule in a fiat regime.

Result 8. The greater the technology of rule, the more difficult the language of rule, if the ruling class is not empty.

All imaginable outcomes can be grouped into three categories: (1) outcomes with only a ruling class, (2) outcomes with only a subject class, and (3) outcomes with both a ruling class and a subject class. Which categories are possible?

An outcome with only a ruling class is impossible. In such an outcome, no wealth would be produced, since only subjects can produce wealth. Rulers, while allocated no wealth, would incur linguistic costs (except at one point on the continuum, where linguistic resistance is 0). Thus, rulers' net wealth rates would be negative. A ruler changing classes and becoming a lone subject would produce wealth at a positive rate. By assumption 2, not all of this wealth would be extracted. Thus, this subject would enjoy a positive net wealth rate. Any ruler would, in this situation, profit by becoming a subject. Therefore, an outcome with only a ruling class is not an equilibrium of class choices and is predicted not to occur. This is result 1.

Outcomes with only a subject class, on the other hand, are possible. We can call these *anarchic* outcomes. If there is only a subject class, the compliance level is 0 and no wealth is extracted. Every subject gets net wealth at a positive rate. If any one subject were to change classes, this lone ruler would, by assumption 3, constitute a ruling class smaller than its minimal size. This means the ruling class would be too small to extract any wealth from the subject class. The lone ruler's net wealth rate would necessarily be zero or negative. Thus, an anarchic outcome would be an equilibrium of class choices. This gives us result 2.

Outcomes with two classes are also possible, as we would expect, and all such outcomes follow a simple pattern. When citizens are arranged along the continuum in order of linguistic resistance, each class consists of a contiguous block of citizens. There is some *critical rate* of linguistic resistance. All citizens with lower rates of linguistic resistance belong to the ruling class, and all citizens with higher rates of linguistic resistance belong to the subject class. Any other outcome would be unstable.

Suppose this pattern were violated, and some ruler had more linguistic resistance than some subject. If the ruler's net wealth rate were the same as or greater than the subject's, the subject could profit by becoming a ruler, because the subject-turned-ruler's net wealth rate would be greater than that of the ruler, hence greater than the subject's current rate. If instead the ruler's net wealth rate were smaller than the subject's, the ruler could profit by becoming a subject. Thus, the pattern of rulers below some critical rate of linguistic resistance, and subjects above that rate, cannot be violated. This is result 3.

The argument for result 3 leads to a further conclusion: the ruling class is privileged. All subjects have the same rate of net wealth, while rulers' net wealth rates vary according to linguistic resistance. But the least wealthy ruler is as wealthy as a subject. We can show this by considering what would happen if it were not true. If any ruler were less wealthy than a subject, such a ruler could simply change classes and rise to the subject's rate of net wealth. And, if the least wealthy ruler were wealthier than a subject, the next citizen on the continuum — differing from this ruler by only an infinitesimal amount of linguistic resistance — could become a ruler and rise to a rate of net wealth only infinitesimally below that of the ruler. An equilibrium of class choices requires the last ruler and all subjects to have the same rate of net wealth. This is result 4.

Result 5 tells us that the ruling class is never larger than its effective size and never smaller than its minimal size, unless it is empty. The impossibility of a smaller-than-minimal ruling class should be no surprise. The net wealth rate of each ruler depends, in general, on the size of the ruling class. However, if the ruling class is smaller than (or equal to) its minimal size, then no wealth is extracted from the subject class (as assumption 3 implies), and yet the rulers incur linguistic costs. Their linguistic costs are proportional to their rates of linguistic resistance, and hence to their positions on the continuum, multiplied by the difficulty of the language of rule. The linguistic cost rate would be 0 for the lone ruler at the beginning of the continuum and positive for all other rulers. No ruler's rate of net wealth would be positive if the ruling class were smaller than, or equal to, its minimal size. The subjects, by contrast, would all enjoy positive rates of net wealth, since none of their wealth would be extracted. Any ruler in this situation could obtain an increased rate of net wealth by becoming a subject. Therefore, an outcome with a ruling class smaller than its minimal size would not be an equilibrium of class choices.

Why, then, is it impossible for the ruling class to be larger than its effective size? This part of result 5 is based on the principle that the outcome must be an equilibrium of voting results. If the ruling class were ever larger than its effective size, it would vote to increase the difficulty of the language of rule, and it would continue doing so until it ceased being larger than its effective size.

I shall introduce a proof of this part of result 5 by giving an example that satisfies the assumptions of the model. Any specific polity can be completely defined in this model by four parameters: the size of the polity, the minimal size of the ruling class, the effective size of the ruling class, and the technology of rule. All else follows from these. The size of the polity is arbitrary: we can choose any positive number for it without affecting the results we get.

In this illustration I shall define the polity as follows:

Polity size	1.00
Minimal ruling-class size	0.05
Effective ruling-class size	0.40
Technology of rule	0.80

In other words, the ruling class needs to grow to 5 percent of the population before it can begin extracting wealth produced by the subject class, and once the ruling class reaches 40 percent of the population it extracts 80 percent of the subject class's wealth, which is the most that can be extracted. Figure 1 shows the extracted proportion of the subject class's wealth (the compliance level multiplied by the technology of rule) as a function of the size of the ruling class.

From this relationship we can derive the rate of net wealth enjoyed by the subjects. Subjects produce wealth at a rate of 1, and they keep whatever isn't extracted. There is no other component in their net wealth rate, and the rate is the same for all subjects. So, by subtracting the curve in Figure 1 from 1 at each point, we obtain the subjects' rate of net wealth, as shown in Figure 2.

Having determined the subjects' rate of net wealth for each possible ruling-class size, we must do the same for the rulers. This is more complicated, because rulers do not share a single rate and because their net wealth does not depend merely on the size of the ruling class. So we can proceed in steps.

Step 1 is to determine the *total produced wealth*, which I have also called the *tax base*. This amount declines steadily with an expanding ruling class, because whenever part of the citizenry moves from the

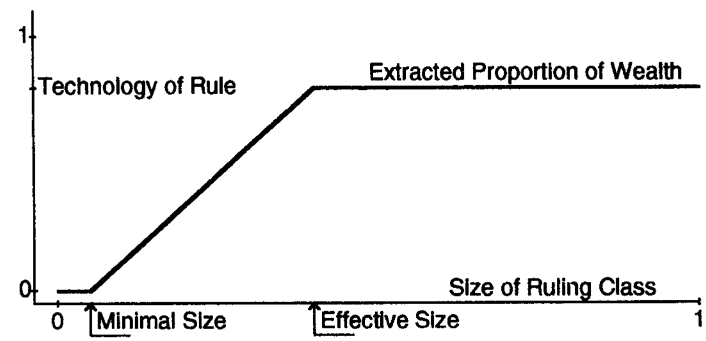

Figure 1. *Extracted wealth as a proportion of total produced wealth*

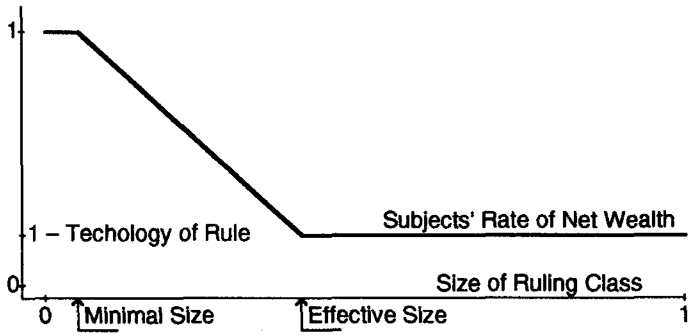

Figure 2. *Subjects' rate of net wealth*

subject class to the ruling class the producing sector of the population is depleted. Figure 3 shows this decline for the same example.

Step 2 is to determine the *total wealth extracted* by the ruling class. This is the total produced wealth, multiplied by the extracted proportion. In other words, at each point we multiply the heights of the curves of Figures 1 and 3. The resulting curve for total extracted wealth is shown in Figure 4.

Step 3 is to obtain the rate at which this total *extracted wealth is allocated* to the individual rulers. This rate tends to rise and fall with the total extracted wealth, but it also falls as the ruling class expands, forcing the extracted wealth to be divided ever more thinly among the rulers. If we divide the height of the curve in Figure 4 by the size of the ruling class at each point, we get the individual allocation rate, shown in Figure 5.

Step 4 is to find the rate of *net wealth* of each ruler. This is the allocation rate minus the linguistic cost rate. The allocation rate depends

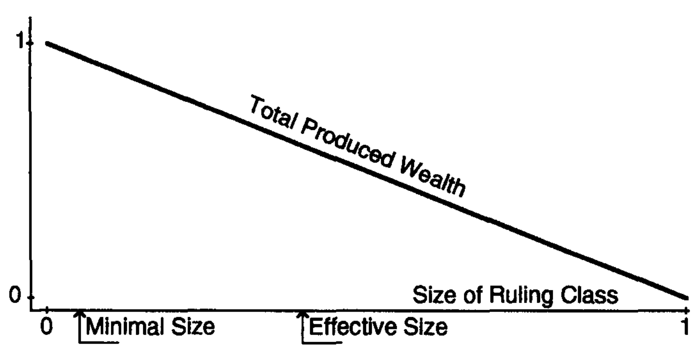

Figure 3. *Total produced wealth*

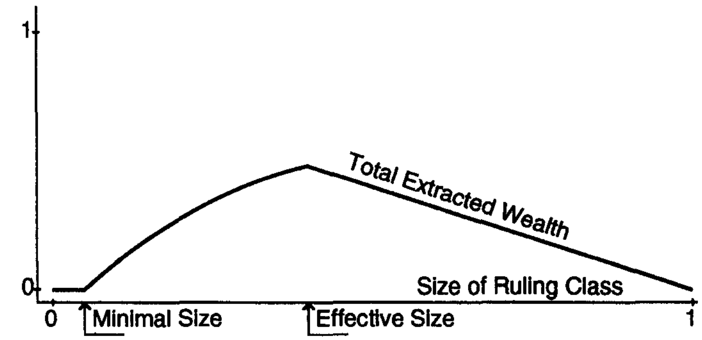

Figure 4. *Total extracted wealth*

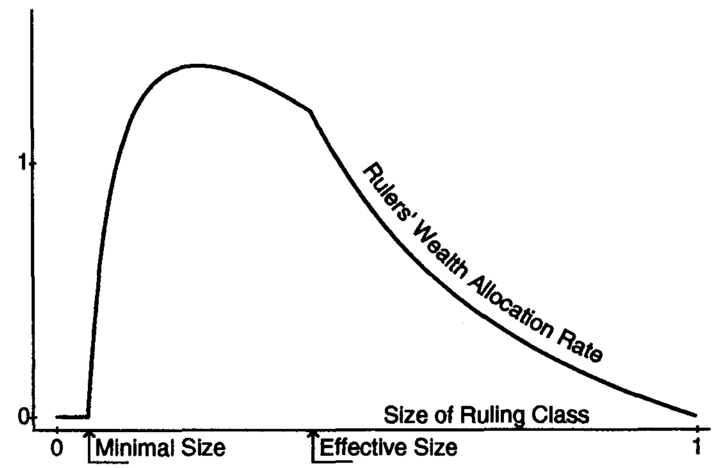

Figure 5. *Rulers' individual rate of allocation of extracted wealth*

on the size of the ruling class, as illustrated by Figure 5. The linguistic cost rate, however, is the product of the difficulty of the language of rule and the rate of linguistic resistance. Linguistic resistance varies from ruler to ruler, growing from left to right on the continuum. Hence, the net wealth rate cannot be graphed simply by lowering the curve in Figure 5. The net wealth curve would depend on how difficult the ruling class has chosen to make the language of rule, and also on the position of the ruler. In a graph I can only give some examples for various combinations of conditions. Figure 6 shows each ruler's net wealth rate for several possible sizes of the ruling class and levels of difficulty of the language of rule. The curves are all linear and downward-sloping from left to right,

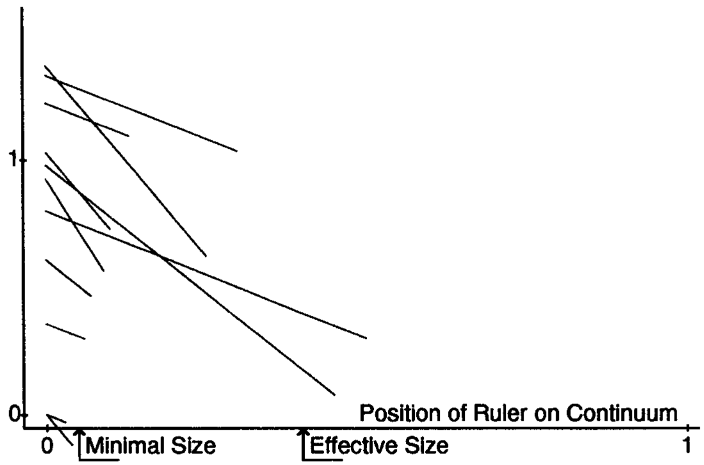

Figure 6. *Rulers' individual rates of net wealth under various conditions*

reflecting the assumption that rulers with higher positions on the continuum have more linguistic resistance and thus pay higher linguistic costs, with linguistic resistance distributed uniformly between 0 and its maximum rate. The right end of each curve has a position on the continuum that corresponds to the size of the ruling class. The slope of each curve reflects the difficulty of the language of rule: the more difficult it is, the steeper the curve.

From the infinity of imaginable outcomes like those represented in Figure 6, we want to know which ones can actually occur. Result 5 says that the ruling class, if it exists, must be between its minimal and effective sizes. So, curves stopping short of the minimal size or extending beyond the effective size cannot occur.

I have already shown that outcomes in which the ruling class is smaller than its minimal size are impossible, but the question remains whether it is possible for the ruling class to be larger than its effective size. Result 5 says that this is impossible. To show this, I shall express the net wealth rates of rulers in a larger-than-effective ruling class as a formula and demonstrate that the formula contradicts itself.

To prepare a formula for the net wealth rates of rulers, I shall adopt convenient scales. I shall set the size of the polity at 1, as in the example on which the graphs are based. I shall also set the maximum rate of linguistic resistance at 1. That will make every citizen's rate of linguistic resistance identical to the citizen's position on the continuum.

I shall use these terms in the formula:

t technology of rule;
m minimal size of the ruling class;
e effective size of the ruling class;
d difficulty of the language of rule;
r ruling-class size;
x ruler at position x on the continuum;
W_x net wealth of ruler x.

If the ruling class were larger than its effective size, the compliance level would be 1, and the total extracted wealth would be equal to the technology of rule, t, multiplied by the total produced wealth, $1 - r$ (the size of the subject class). This total extracted wealth, $(1 - r)t$, would be allocated uniformly to all rulers, making the rate of allocation to each ruler $\frac{(1-r)t}{r}$. Each ruler's rate of linguistic resistance would be equal to that ruler's position on the continuum, x. Therefore, each ruler's linguistic cost rate, the product of linguistic resistance and linguistic difficulty, would be xd.

The net wealth rate of any ruler is the rate of wealth allocation, reduced by that ruler's rate of linguistic cost. Combining the terms in the previous paragraph, we get this formula when the ruling class is larger than its effective size:

$$(1) \quad W_x = \frac{(1-r)t}{r} - xd = \left(\frac{1}{r} - 1\right)t - xd.$$

To prove that equation (1) is contradictory, I shall use result 4, which has already been proved. Result 4 says that the subject's rate of net wealth is the same as that of the last ruler. To see what this result implies, look back at Figure 6. The lower-right ends of the various curves show the position of the last ruler on the continuum and that ruler's rate of net wealth. Now look at Figure 2. It shows the subjects' rate of net wealth. What result 4 tells us is that the lower-right end of any curve in Figure 6 must exactly coincide with the curve in Figure 2. Any curve in Figure 6 that fails that test is predicted not to occur. If the conditions generating such a curve were to exist, either some subjects would choose to become rulers or some rulers would choose to become subjects. An equilibrium of class choices would not be present.

If the ruling class were larger than its effective size, the subjects' rate of net wealth would be $1 - t$, as Figure 2 illustrates. Thus, the last ruler's rate of net wealth would also be $1 - t$. Equation (1) gives us another expression for the last ruler's rate of net wealth:

$$(2) \quad W_r = \left(\frac{1}{r} - 1\right)t - rd = \frac{t}{r} - t - rd.$$

Combining these requirements, we find that

$$(3) \quad \frac{t}{r} - t - rd = 1 - t$$

or, equivalently,

$$(4) \quad d = \frac{t}{r^2} - \frac{1}{r} = \frac{t-r}{r^2}.$$

Equation (4) gives a relationship that must exist between the size of the ruling class and the difficulty of the language of rule, if the ruling class is larger than its effective size. This relationship is based on result 4, which in turn is derived from the requirement that the outcome be an equilibrium of class choices. Consider the example I have used above, where $e = 0.4$. Suppose the ruling class wanted to recruit enough members to bring its size to 0.5, which is above the effective size. Equation (4) tells the ruling class what it needs to do. Substituting 0.5 for r in the equation, the ruling class would learn that it must give the language of rule a difficulty of 1.2.

So far, I have not shown that a ruling class larger than e is contradictory. But I have used only one of the two equilibrium principles: the principle of an equilibrium of class choices. I have not yet used the principle of an equilibrium of voting results. This principle imposes further restrictions, and it is the two equilibrium principles in combination that turn out to contradict the assumption that the ruling class is larger than its effective size.

To see how the contradiction emerges, suppose the rulers in our example have given the language of rule a difficulty of 1.2 and the ruling class now has a size of 0.5. When the rulers next consider whether to change the difficulty of the language of rule, how will they vote? In general, rulers at low positions on the continuum (to the left side of Figure 6) will tend to vote for a more difficult language, because their learning costs are low and the increased linguistic costs will be small for them compared with their increased allocations of extracted wealth. Those with high positions (to the right in the figure) will tend to vote for an easier language, because their linguistic costs are higher. Somewhere there may be a ruler whose learning cost and allocation rates would change by identical amounts and who would therefore be indifferent between the status quo and a proposed change in either direction. If that person is located to the left of the halfway point within the ruling class, a majority will vote for an easier language. If this ruler is to the right of the halfway point, a majority will vote for a more difficult language. Only if the ruler is precisely at the halfway point will neither proposal be adopted. Only then is the outcome an equilibrium of voting results.

To predict the result of a vote in this example, examine Figure 7. It graphs the rates of net ruler wealth for three different levels of difficulty of the language of rule, all of which would produce a ruling class larger than the e. (The middle curve represents the 50 percent-rulers outcome discussed above.)

If the rulers make the language slightly more difficult, the net wealth curve becomes steeper. (The new line is steeper because a more difficult language makes differences in linguistic resistance count more heavily in the rulers' net wealth rates.) Approximately the left $\frac{2}{3}$ of the rulers' net wealth rates rise, and approximately the right $\frac{1}{3}$ of the rulers' net wealth rates fall. All rulers situated to the left of the pivot around which the thin curves appear to turn, gain from a contraction in the ruling class. They vote for a more difficult language of rule and win the vote. This shows that the ruling class, while it could make itself larger than its effective size, would not do so. An outcome with a difficulty of 1.2 and a ruling-class size of 0.5 satisfies the class-choice equilibrium test, but it fails the voting equilibrium test.

What we just saw in this particular case holds in all cases. When the ruling class is larger than its effective size, the median ruler (the ruler halfway between 0 and r) can't be the pivotal voter.

The proof can begin with a reformulation of equation (1). That equation contains d. Equation (4) gives us a formula for d. If we substitute that formula for d in equation (1), it becomes

$$(5) \quad W_x = \left(\frac{1}{r} - 1\right)t - x\frac{t-r}{r^2} = \frac{x+t}{r} - \frac{xt}{r^2} - t.$$

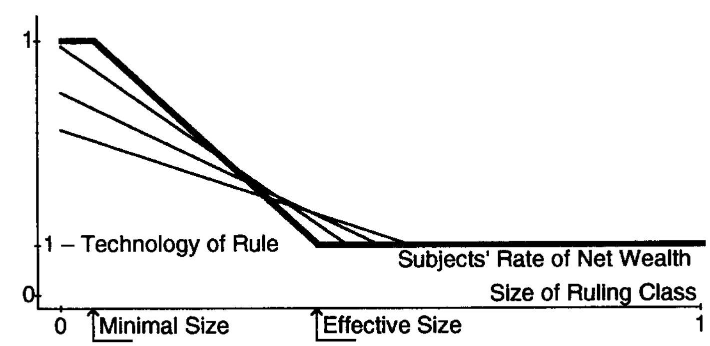

Figure 7. *Failure of a class equilibrium to be a voting equilibrium*

Any ruler x is pivotal if the rate at which W_x changes when r begins to change is 0. This rate of change of W_x can be computed according to the rules of derivative calculus, assuming (as I shall here) that W_x is smooth. The change of W_x per positive unit change of r is

$$(6)\quad \frac{\partial W_x}{\partial r} = \frac{2xt}{r^3} - \frac{x+t}{r^2}.$$

If x is pivotal then this expression equals 0, and we can multiply it by r^2 and it will still equal 0. Consequently,

$$(7)\quad \frac{2xt}{r} - x - t = 0$$

defines the pivotal voter among the rulers. We have an equilibrium of voting results only when its ruler x is the median ruler, that is, when $x = \frac{r}{2}$. So, let us replace x with $\frac{r}{2}$ in equation (7). That replacement gives us

$$(8)\quad 0 = \frac{\frac{2r}{2}t}{r} - \frac{r}{2} - t = t - \frac{r}{2} - t = -\frac{r}{2}.$$

But this equation can be true only if $r = 0$, that is, if there is no ruling class at all.

Thus, the supposition that the ruling class is larger than its effective size has yielded a conclusion that contradicts that supposition. If $r > e$ and $e > 0$, then r cannot be 0. This completes the proof of result 5.

It is now clear that, if there is a ruling class at all, it is somewhere between its minimal size and its effective size. If it were smaller than its minimal size its members would be suffering from negative rates of net wealth and would abandon their class. And I have just proved that if it were larger than its effective size the rulers would vote to change the difficulty of the language of rule.

Exactly how big will the ruling class be, then? The principle of an equilibrium of class choices, combined with the principle of an equilibrium of voting results, narrows the predicted outcomes to a unique one.

If the ruling class is between its minimal size and its effective size, then the compliance level can be formulated as $\frac{r-m}{e-m}$. With this formula, when the ruling class is at its minimal size ($r = m$) the compliance level is 0, and when the ruling class reaches its effective size ($r = e$) the compliance level reaches 1. The proportion of the produced wealth that the rulers extract, shown in Figure 1, is the product of this compliance level and the technology of rule: $\frac{r-m}{e-m}t$. The total produced wealth, shown in Figure 3, is $1 - r$. The total extracted wealth, shown in Figure 4, is the product of these last two amounts: $(1-r)\frac{r-m}{e-m}t$. The allocation per ruler, shown in Figure 5, is this total divided by r. When we finally subtract a ruler's rate of linguistic cost, we have the ruler's rate of net wealth:

$$(9)\quad W_x = \frac{(1-r)\frac{r-m}{e-m}t}{r} - xd = \left(\frac{1}{r} - 1\right)\frac{r-m}{e-m}t - xd.$$

We must now force this equation to conform to the principles of an equilibrium of class choices and an equilibrium of voting results. The size of the ruling class will follow from these requirements.

In an equilibrium of class choices, the last ruler's (ruler r's) rate of net wealth is the same as any subject's. A subject's net wealth rate is 1 minus the extracted proportion of produced wealth. Thus, we obtain this formulation of the equilibrium of class choices:

$$(10)\quad 1 - \frac{r-m}{e-m}t = W_r = \left(\frac{1}{r} - 1\right)\frac{r-m}{e-m}t - rd.$$

It can be simplified to

$$(11)\quad rd = \frac{t}{r} \times \frac{r-m}{e-m} - 1.$$

To narrow the possible outcomes to a single prediction, we need to determine which of the outcomes that equation (11) allows would survive a vote of the ruling class. We begin deriving this equilibrium of voting results by using equation (11) to express d in terms of r:

$$(12)\quad d = \frac{t}{r^2} \times \frac{r-m}{e-m} - \frac{1}{r} = \left(\frac{t}{e-m} - 1\right)\frac{1}{r} - \frac{tm}{(e-m)r^2}.$$

Then we substitute this expression for d into equation (9):

$$(13)\quad W_x = \left(\frac{1}{r} - 1\right)\frac{r-m}{e-m}t - \left(\frac{t}{e-m} - 1\right)\frac{x}{r} + \frac{tmx}{(e-m)r^2}$$
$$= \frac{t}{e-m}\left[1 - \frac{m}{r} - r + m - \frac{x}{r} + \frac{(e-m)x}{tr} + \frac{mx}{r^2}\right].$$

For an outcome to be a voting equilibrium, the median ruler must not profit from a slight change in the size of the ruling class. The rate of change of W_x when r changes must be 0 when $x = \frac{r}{2}$. This rate of change, according to rules of differentiation, is

$$(14)\quad \frac{\partial W_x}{\partial r} = \frac{t}{e-m}\left[\frac{m}{r^2} - 1 + \frac{x}{r^2} - \frac{(e-m)x}{tr^2} - \frac{2mx}{r^3}\right].$$

If this expression is to equal 0 when $x = \frac{r}{2}$, we get the following requirement:

$$(15)\quad 0 = \frac{t}{e-m}\left[\frac{m}{r^2} - 1 + \frac{1}{2r} - \frac{e-m}{2tr} - \frac{m}{r^2}\right] = \frac{t}{e-m}\left[\frac{1}{2r} - \frac{e-m}{2tr} - 1\right]$$
$$= \frac{t-e+m-2tr}{2(e-m)r}.$$

Equation (15) implies that

$$(16)\quad 2tr + e - m - t = 0,$$

which in turn implies that

$$(17)\quad r = \frac{1}{2} - \frac{e-m}{2t}$$

is the formula for the exact size of the ruling class. Inspecting this formula, we can see some constraints on the ruling class. The right-hand fraction of this expression must be positive, since e (the effective size) is greater than m (the minimal size). Consequently, the ruling class must be smaller than $\frac{1}{2}$. In other words, it must constitute a minority of the population, if it exists at all. With equation (17), then, result 6 is proved.

Applied to our example, equation (17) predicts that the ruling class will constitute about 28 percent of the population. When this outcome is graphed, we can see that the pivotal ruler is indeed at the median position within the ruling class. Figure 8 shows the distribution of net wealth rates among the rulers if they set the difficulty of the language of rule slightly above or slightly below the predicted level. Any slight change at this level benefits half the rulers and hurts the other half, yielding no majority for a change in either direction.

Where the polity of our example ruled according to a fiat regime instead of a linguistic regime, the ruling class would make up only about

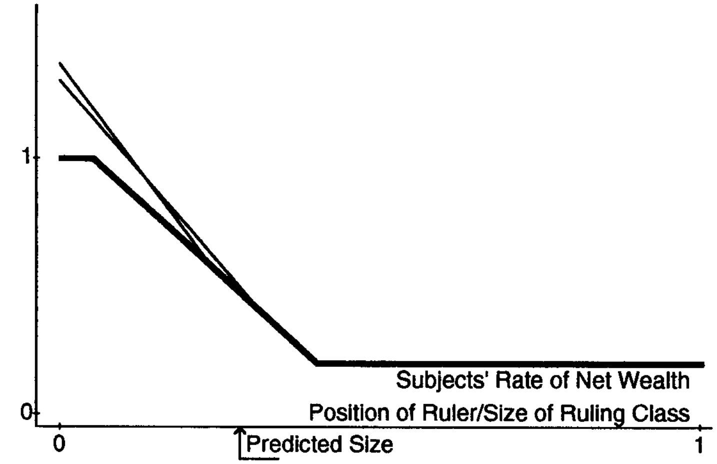

Figure 8. *Predicted outcome*

22 percent of the population, leaving the subjects with more net wealth than they enjoy under the linguistic regime. The figure of 22 percent is the position where the wealth allocation rate, graphed above in Figure 5, reaches its peak. By experimentation with various values of m, e, and t, one can also find cases where the difference is opposite, with the ruling class smaller in a linguistic regime than it is in a fiat regime. But this example shows definitively (for my model) that the speculation I offered earlier is wrong: a linguistic regime does not always guarantee any reduction in the power of the rulers to extract their subjects' wealth.

Another conclusion emerging directly from equation (17) is that the ruling class varies in size directly with the technology of rule. As t rises, the right-hand side of equation (17) gains in value. If we surmised that an increase in the efficiency of wealth extraction might lead to a smaller ruling class, we would be wrong. As its maximum extractive power rises, the ruling class only grows larger. In an equilibrium model, this should not be startling. Citizens decide individually which class to join, and when the technology of rule rises, membership in the ruling class becomes more advantageous, while membership in the subject class becomes less advantageous.

What happens in a fiat regime, where the ruling class simply decides on its size? There the whole ruling class gains or loses together, except for the infinitesimal fraction of rulers who join or leave the class during a change. There is a uniform rate of net wealth, consisting only of the rate of allocation of extracted wealth.

This rate is shown in equation (9), except that we omit the term for linguistic cost:

$$(18)\quad W = \left(\frac{1}{r} - 1\right)\frac{r-m}{e-m}t = \left(1 - \frac{m}{r} - r + m\right)\frac{t}{e-m}.$$

W reaches its maximum when its rate of change with r is 0. Differentiation reveals that this condition is

$$(19)\quad \frac{\partial W}{\partial r} = \frac{t}{e-m}\left(\frac{m}{r^2} - 1\right),$$

which is true only when $m = r^2$.

In a fiat regime, it turns out that the equilibrium size of the ruling class is simply the square root of m, the minimal size of that class. If the minimal size is 0.05, as in the example I have been using, then the ruling class's predicted size is the square root of 0.05, which is, as mentioned above, about 0.22.

The size of the ruling class, then, is completely independent of the technology of rule in a fiat regime. The fact that the ruling class expands when the technology of rule increases is peculiar to a linguistic regime and not shared by a fiat regime. This is result 7. Why don't more citizens join the ruling class of a fiat regime when the rulers' wealth is increased by a rise in the technology of rule? Because the ruling class decides, by fiat, to keep them out. And nothing motivates the ruling class in a fiat regime to reconsider its optimal size. Changes in the technology of rule raise or lower the rulers' wealth allocation rate (the curve in Figure 5) proportionally, leaving its peak at a constant horizontal position.

The expansion of the ruling class as a result of an increase in the technology of rule in a linguistic regime does not come about all by itself. The rulers and subjects respond with their class choices and votes in such a way that this effect takes place. As they adjust to an increased technology of rule, we find them also adjusting to difficulty of the language of rule.

The result is to make the language of rule more difficult, and this, which is result 8, may be a surprise. After all, if the ruling class wants to become *smaller* it makes the language of rule more difficult. So, how can an increase in difficulty accompany an *enlarged* ruling class? The reason is that two things are changing at once: the difficulty of the language of rule and the technology of rule. The increased technology of rule raises the net wealth of all rulers and makes more citizens want to be rulers. To limit the influx of new rulers, the ruling class's majority must make its language more difficult.

For a proof of result 8, we can convert equation (17) into an expression for the technology of rule,

$$(20)\quad t = \frac{e-m}{1-2r},$$

and then substitute this expression for t into equation 12:

$$(21)\quad d = \frac{t}{r^2} \times \frac{r-m}{e-m} - \frac{1}{r} = \frac{\frac{e-m}{1-2r}}{r^2} \times \frac{r-m}{e-m} - \frac{1}{r} = \frac{r-m}{(1-2r)r^2} - \frac{1}{r}$$

$$= \frac{r-m-(1-2r)r}{(1-2r)r^2} = \frac{2r^2-m}{(1-2r)r^2} = \frac{2-\frac{m}{r^2}}{1-2r}.$$

An increase in r makes the numerator of the final expression in equation (21) increase, and it makes the denominator decrease. The value of d inevitably rises. We know that when the technology of rule rises the ruling class must expand, and now equation (21) tells us that when the ruling class expands the difficulty of the language of rule must also increase. This proves result 8.

When two polities differ in the difficulty of their languages of rule, what can we conclude about them? If their minimal and effective ruling-class sizes are the same, then the difference between their languages of rule can only be due to a difference in technologies of rule. The polity with the higher technology of rule must then have a more difficult language of rule, a larger ruling class, and a poorer subject class. If the language is more difficult and the ruling class is larger, then (as the discussion of Figure 6 should make plain) the gap between the richest ruler and the poorest ruler must also be greater.

Thus, a language of rule is an indicator of two kinds of inequality. (1) The more difficult the language of rule, the less the subject class retains of its per-capita wealth. (2) The more difficult the language of rule, the greater the gap between the net wealth rates of any two rulers. A difficult language of rule creates inequality not only between rulers and subjects, but also among rulers. It makes some forgo the benefits of power, while it makes those who claim power pay unequal costs.

Why do rulers voluntarily make the language of rule difficult and incur the cost of mastering it? The only reason is that their status as rulers is lucrative, and they need to make the language difficult in order to keep subjects from joining the rulers, since a larger ruling class would reduce the polity's total product and dilute the rulers' per-capita gains. Thus, a difficult language of rule tells us that the popular pressure on the ruling class for a share of the spoils must be great. And, heavy pressure to enter the ruling class tells us that rulers must be enjoying great premiums in net wealth over the subject class.

Seeing this dual inequality, what changes would egalitarians demand? Would they agitate for the simplification of the language of rule, or even for its removal as the criterion of ruling-class membership? How would an egalitarian with dictatorial power change the difficulty of the language of rule, if that were the only lever with which to reform the polity? Would an egalitarian give the language a difficulty of 0, making it identical to the popular native language, so that everyone mastered it as a costless byproduct of growing up?

The answer (this time without proof): an egalitarian would seek to make the language of rule *more* difficult, not less. By reviewing the foregoing model, we can see intuitively that there are at most two ways to achieve an egalitarian outcome, namely an outcome in which every citizen obtains net wealth at the same rate. One can set the difficulty of the language of rule at 0. The result will be a large ruling class (if any) and an equal but low rate of net wealth. In the example I have been using, the ruling class would expand to 80 percent of the population, and everyone's net wealth rate would sink to 0.2 (subjects' originally predicted net wealth rate is about 0.47). Alternatively, one can set the difficulty of the language of rule high enough to deter everyone from becoming a ruler. In my example, the predicted difficulty of the language of rule set by the ruling class is about 3.1. An egalitarian dictator would need to raise this to about 3.6 or more, and then the ruling class would disappear. All citizens would have identical net wealth rates, but this time their rate would be at its maximum possible level, 1. A difficult language of rule may cause inequality, but the most effective means of eliminating the inequality may be to make the language even more difficult than it already is.

University of Washington

References

Das Gupta, Jyotirindra; and Gumperz, John J. (1968). Language, communication and control in North India. In *Language Problems of Developing Nations*, Joshua A. Fishman, Charles A. Ferguson, and Jyotirindra Das Gupta, (eds.), 151–166. New York: Wiley.

Durkheim, Emile (1964 [1893]). *The Division of Labor in Society*. New York: Macmillan.

Edelman, Murray (1977). *Political Language: Words that Succeed and Policies that Fail*. New York: Academic Press.

Haugen, Einar (1966). *Language Conflict and Language Planning: The Case of Modern Norwegian*. Cambridge, MA: Harvard University Press.

Jespersen, Otto (1964 [1925]). *Mankind, Nation and Individual from a Linguistic Point of View*. Bloomington: Indiana University Press.

Levi, Margaret (1983). The predatory theory of rule. In *The Microfoundations of Macrosociology*, Michael Hechter (ed.), 216–249, Philadelphia: Temple University Press.

Lindblom, Charles E. (1977). *Politics and Markets*. New York: Basic Books.

Myers-Scotton, Carol (1989). Elite closure as boundary maintenance: the case of Africa. Unpublished manuscript.

Selten, Reinhard; and Pool, Jonathan (1982). Ĉu Mi Lernu Esperanton? Enkonduko en la Teorion de Lingvaj Ludoj [Symmetric Equilibria in Linguistic Games]. Working Papers No. 112, Institute of Mathematical Economics, University of Bielefeld.