# Adam Smith's Rational Choice Linguistics

## Language and Other Human Actions

Linguistics may be the last of the social sciences to avoid the rational choice paradigm. Many philosophers, Adam Smith especially, argue that being human is the same as using language. Reason and speech are primitives for Smith; we no more choose to use language than we choose to be human. His argument in the *Wealth of Nations* is that trade and language are two aspects of the same process. Humans trade because we have language; nonhumans do not trade because they do not.[1] Merely because language is a background condition for human choices, however, does not obviate for Smith rational choice *aspects* of language.

Smith's argument exploits the simple possibility of substitution of one feature of language for another in such a way as to minimize the time cost of conveying meaning. He distinguishes sharply between what is true for children naturally learning a language and what is true for adults learning a language in which to trade. Perhaps because the rational choice basis of such trade languages (pidgins) is entirely obvious, pidgins have been *defined* out of existence to keep traditional linguistic theory safe from embarrassing contact with rational choice considerations.[2]

The claim that a language is spoken by those of homogeneous competence is regarded as true by definition. This web of belief traps rational choice considerations because now, by definition, we have precluded the possibility that

---

1. Smith's argument is discussed in Levy 1992 in light of modern experimental economics, which finds that (1) rats have preferences but (2) rats don't trade. Steven Pinker (1994, 340–41) discusses continuing attempts to teach chimpanzees sign language. He does not recognize the importance of the experimental result that chimpanzees who can sign will cooperate more.

2. Here is a Chomskian simply defining pidgin out of existence as a "language" spoken by humans: "Thus we cannot say things like *Last night I slept bad dreams a hangover snoring no pajamas sheets were wrinkled,* even though a listener could guess what that would mean. This marks a major difference between human languages [*sic*] and, for example, pidgins and the signing of chimpanzees, where any word can pretty much go anywhere" (Pinker 1994, 117).

one's fluency in a language increases with exposure.[3] Scholars working in pidgin and/or language death are leading the way in questioning the definitional status of such a specification.[4] Why this might be so will be clearer after we look at Smith's account of a language that loses its grammar.

## In Smith's Words

In standard histories of linguistics such as Land 1974, Smith is given credit for developing the argument that languages will grow more grammatically complex over time. It is insufficiently noticed that Smith restricts this argument to an isolated language. His example is classical Greek, a language in which he believed all the words were generated internally.[5] The Smithian problem upon which we shall focus is how the grammar of language evolves as the result of trade between people with different languages.

Smith does not see any reason why a language learned in childhood could not have an arbitrarily large number of what modern linguistics calls inflections:

> As long as any language was spoke by those only who learned it in their infancy, the intricacy of its declensions and conjugations could occasion no great embarrassment. The far greater part of those who had occasion to speak it, had acquired it at so very early a period of their lives, so insensibly and by such slow degrees, that they were scarce ever sensible of the difficulty. (*Rhetoric,* 220)

Something very interesting happens when adults need to conduct business across languages. Rational choice considerations enter into the choice of a language's grammatical structure:

---

3. "What we say is that the child or foreigner has a 'partial knowledge of English,' or is 'on his or her way' toward acquiring knowledge of English, and if they reach the goal, they will then know English. Whether or not a coherent account can be given of this aspect of the commonsense terminology, it does not seem to be one that has any role in an eventual science of language" (Chomsky 1986, 16).

4. David DeCamp (1977, 9) tells of the first great student of pidgins and Creoles: "At one time he was warned by a senior colleague that he should abandon this foolish study of funny dialects and work on Old French if he wished to further his academic career. [Note:] History indeed repeats itself. When I myself began studying Jamaican Creole in 1957, I received from a colleague a similar warning that I should avoid such quasi-languages and should work on an American Indian or other 'real' language." "Pidgins and creoles were long the neglected step-children of linguistics because they were thought to be marginal, and not 'real' full-fledged languages" (Romaine 1988, 1).

5. "The Greek seems to be, in a great measure, a simple, uncompounded language, formed from the primitive jargon of those wandering savages, the ancient Hellenians and Pelasgians, from whom the Greek nation is said to have been descended. All the words in the Greek language are derived from about three hundred primitives, a plain evidence that the Greeks formed their language almost entirely among themselves, and that when they had occasion for a new word, they were not accustomed, as we are, to borrow it from some foreign language, but to form it, either by composition, or derivation from some other word or words, in their own. The declensions and conjugations, therefore, of the Greek are much more complex than those of any other European language with which I am acquainted" (*Rhetoric,* 222).

But when two nations came to be mixed with one another, either by conquest or migration, the case would be very different. Each nation, in order to make itself intelligible to those with whom it was under the necessity of conversing, would be obliged to learn the language of the other. The greater part of individuals too, learning the new language, not by art, or by remounting to its rudiments and first principles, but by rote, and by what they commonly heard in conversation, would be extremely perplexed by the intricacy of its declensions and conjugations. They would endeavor, therefore, to supply their ignorance of these, by whatever shift the language could afford them. Their ignorance of the declensions they would naturally supply by the use of prepositions; . . . The same alteration has, I am informed, been produced upon the Greek language, since the taking of Constantinople by the Turks. *The words are, in a great measure, the same as before; but the grammar is entirely lost,* prepositions having come in the place of the old declensions. This change is undoubtedly a simplification of the language, in point of rudiments and principle. It introduces, instead of a great variety of declensions, one universal declension, which is the same in every word, of whatever gender, number, or termination. (*Rhetoric,* 220–21; emphasis added)

The phrase emphasized in this passage—"the grammar is entirely lost"—is a signature of a type of language heavily studied by modern linguists where *degrammaticalization* is a term of choice (Romaine 1989, 379). Continuing with Smith, we read:

A similar expedient enables men, in the situation above mentioned, to get rid of almost the whole intricacy of their conjugations. There is in every language a verb, known by the name of the substantive verb; in Latin, *sum;* in English, *I am.* This verb denotes not the existence of any particular event, but existence in general. It is, upon this account, the most abstract and metaphysical of all verbs; and, consequently, could by no means be a word of early invention. When it came to be invented, however, as it had all the tenses and modes of any other verb, by being joined with the passive participle, it was capable of supplying the place of the whole passive voice, and of rendering this part of their conjugations as simple and uniform, as the use of prepositions had rendered their declensions. (*Rhetoric,* 221)[6]

---

6. Smith writes for an audience with a background in Latin grammar. When Latin teachers today explain the declension of nouns and the agreement of adjectives to English speakers, they appeal to English prepositions. "Consider the Latin roots *port* (gate) and *magna* (large). The genitive case, *portae magnae,* is 'of' the large gate, and so on for cases which cover 'to/for,' 'by/with/from,' etc." (Wheelock 1992, 7–8).

## A Model of Inflection and Trade

Our jumping off point is Smith's claim that language is like a machine.[7] This encourages us to think about language in terms of production functions. In particular, let us think about the production of meaning by different properties of language. Inflections can be viewed as a method of economizing on vocabulary. Knowledge of the root and inflectional formula—this is what linguistics call a "paradigm"—allows one to solve for the right word. To the language learner, a grammatical paradigm is really closer in spirit to a regression equation than to a nonstochastic equation because of grammatical irregularity. Irregularity is the paradigm's residual.

Consider the English inflection for number with its two explicit cases: singular and plural. The paradigm is to add an *s* if it is plural and do nothing to the root if it is singular:

cow + *s* (if more than one cow)

cow (if one cow).

There are wonderfully bizarre exceptions, or residuals in regression terminology, to the paradigm. *Child* changes to *children, goose* changes to *geese, mouse* changes to *mice, wharf* changes to *wharves,* and *dwarf* changes either to *dwarves* or *dwarfs* depending upon whom you read. Steven Pinker (1994, 141–43) claims that these irregularities are themselves the vestigial evidence of other rules for inflecting number. If this is so, then the residuals come from something akin to random regime shifts.

Do we need this explicit inflection? English speakers seem not to be terribly disadvantaged by the lack of an explicit marker for the dual case. Surely, we could treat *cow* the way we treat *deer.* There is exactly the same word for one deer as for many deer; we let plurality be indicated with a cardinal number.

---

7. "It is in this manner that language becomes more simple in its rudiments and principles, just in proportion as it grows more complex in its composition, and the same thing has happened in it, which commonly happens with regard to mechanical engines. All machines are generally, when first invented, extremely complex in their principles, and there is often a particular principle of motion for every particular movement which it is intended they should perform. Succeeding improvers observe, that one principle may be so applied as to produce several of those movements; and thus the machine becomes gradually more and more simple, and produces its effects with fewer wheels, and fewer principles of motion. In language, in the same manner, every case of every noun, and every tense of every verb, was originally expressed by a particular distinct word, which served for this purpose and for no other. But succeeding observation discovered, that one set of words was capable of supplying the place of all that infinite number, and that four or five prepositions, and half a dozen auxiliary verbs, were capable of answering the end of all the declensions, and of all the conjugations in the ancient languages" (*Rhetoric,* 223–24).

Alternatively, we could have a special word for a multitude of cows; indeed, we have one, *cattle*.[8]

Let us consider other inflections. Suppose we wish to convey information that a certain subject, Robert, beat a certain object, John. Different languages have different methods for this. Inflected languages may have a nominative inflection that has a marker indicating what part of speech the noun fulfills. Once the markers are in place, the meaning is fixed whether we write the subject, the object, or the verb first. Word order may be selected on poetic grounds.[9] In English, except for pronouns, inflections indicating differences between subject and object do not exist.[10]

The alternative technology for indicating subject and object is fixed word order. Comparing English to inflected classical languages, such as Greek and Latin, Smith advises writers to accept the natural structure of English (*Rhetoric*, 225). It turns out that the advice is efficient (Diamond and Levy 1994).

The consideration that vocabulary and inflections are alternative methods for conveying information encourages us to us write the production of meaning as a function of the size of vocabulary and the complexity of the inflections. The number of inflections in the language could measure the inflectional dimensionality of the language.

Meaning = *M*(Vocabulary, Inflections).

We suppose that we can extend the visual representations of equal quantities of output—isoquants—to draw isomeaning relationships for any language. Figure 5 gives an equal isomeaning relationship in each of two languages, a vocabulary-rich E and an inflection-rich G. We suppose a time budget for children (*cc* in fig. 5) such that the choice of E or G is a matter of indifference. Given the budget *cc,* an equal amount of information can be conveyed with an equal expenditure of time. This specification is a simple consequence of something agreed upon by serious linguistic scholars: all languages learned from childhood are equally good at conveying any sort of information. As Pinker (1994, 27) observes, there may be people with Stone Age technology; however, there are no Stone Age languages.

---

8. The example points to the complicated interrelations between gender and number in English: *Cow* has feminine gender; *cattle* does not. The lack of an ungendered pronoun of singular number is a source of much discomfort in modern style manuals.

9. "The variety of termination in the Greek and Latin, occasioned by their declensions and conjugations, gives a sweetness to their language altogether unknown to ours, and a variety unknown to any other modern language" (*Rhetoric,* 224).

10. It is interesting that the pair *who* and *whom* are losing their distinction in standard American English (Pinker 1994, 116). In radio American English, other pairs of pronouns are merging.
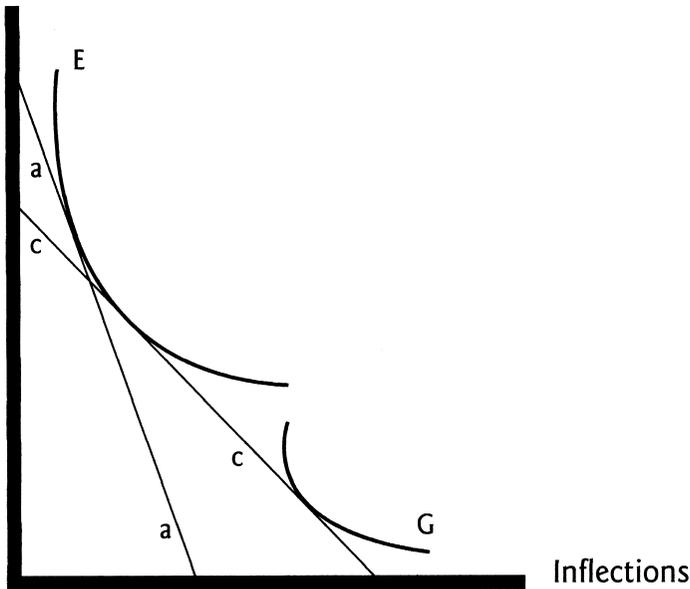
## Vocabulary



Fig. 5.   The production of meaning

But this is only supposed to be true for languages learned from childhood. Suppose, now, that adults wish to conduct business with this amount of meaning. Which language will they pick? We have agreed that from the point of view of children it would not make any difference. Can we say anything about *adult* budget constraints? Smith asserts that it is relatively more difficult in terms of time for adults to pick up inflection patterns than vocabulary. Without inflections, words are words. The word for *pen* in French is no more problematic than the word for *plume* in English. But knowledge of the inflections of one language might not predict another's inflections.

The reason for this is simply that, while some grammatical distinctions (e.g., number) are simple correlatives of the world outside language, other grammatical distinctions (e.g., gender) are wound up in a cultural web of belief.[11] The inflection of number seems to be linked to the world in a way that

---

11. As Dan Slobin (1993, 247) notes, "I would imagine, for example, that if your language lacked a plural marker, you would not have insurmountable difficulty in learning to mark the category of plurality in a second language, since this concept is evident to the nonlinguistic mind and eye. Or if your language lacked an instrumental marker it should not be difficult to learn to add a grammatical inflexion to nouns that name objects manipulated as instruments. Plurality and manipulation are notions that are obvious to the senses in ways that, for example, definiteness and relative tense are not."

the inflection of gender is not. If "there are two cows out there" is true, then there really are two cows out there. But even if "the *Enterprise* launched an F–18 from her deck" is true, aircraft carriers are not obviously female, even though warships are gendered feminine in demotic English. While standard European languages are gendered on the basis of sex, other languages are gendered on many other principles.[12]

Moreover, there are different ways to slice common experience about movements in time and space. For instance, one language, such as English, has verbs conjugated on the basis of time. Other languages, such as Navaho, have verbs conjugated on the basis of aspect.[13] Linguistics since Smith has discovered other inflections in isolated languages, for example, the ergative inflection, which indicates transitivity and the location markers (Dixon 1989, 99, 162).

The difference in the dimensions of language means that one would have to find the structure of the inflectional equations by a trial and error procedure in which the dimension of the specification search increases with the number of inflections of the language.[14] Thus, the adult language learner confronts the full horror of exploratory data analysis in an unknown number of dimensions (Mosteller and Tukey 1977). Inflections often involve a worldview that children learn simultaneously with the language.[15] The adult language learner has to either replace one worldview with another or switch between them on demand.[16]

If Smith is right, then the adult budget constraint will look something like that of *aa* in figure 5. This suggests that languages that adults learn in which to conduct business will be richer in vocabulary than inflections.[17] Of course, lan-

---

12. R. M. W. Dixon (1989, 77) reports that Dyirbal has four genders: masculine, feminine, neuter, and edible. Pinker (1994, 127) gives other wonderful examples.

13. According to Clyde Kluchkhohn and Martha Leighton (1956, 194): "Aspect defines the geometrical character of an event, stating its definability with regard to line and point rather than its position in an absolute time scale or in time as broken up by the moving present of the speaker. . . . Thus, the momentaneous aspect in Navaho means that action is thought of as beginning and ending within an instant, while the continuative suggests that action lasts. Inceptive, cessative, durative, imperfective, and semelfactive, are some of the other aspects in Navaho—with a different paradigm of every verb stem for each."

14. Dixon (1989, 103) describes the problem: "Just asking from English may reveal similarities and likenesses to English patterns, but it is unlikely to uncover things in a language which *do not occur* in English—for example, the complex set of forms for uphill, downhill, upriver, downriver and across the river."

15. Dixon (ibid., 132–35) reports that gendering in Jirrbal encodes creation beliefs.

16. Kluchkhohn and Leighton (1956, 184) suggest that "Every language is a different system of categorizing and interpreting experience. This system is the more insidious and pervasive because native speakers are so unconscious of it as a system, because to them it is part of the very nature of things, remaining always in the class of background phenomena. . . . They take such ways of thought as much for granted as the air they breathe, and unconsciously assume that all human beings in their right minds think the same way."

17. "The remarkable freedom of English construction is not the least cause of the readiness of foreigners to use it among themselves, in preference to their own less analytical tongues" (Ogden 1935, 108).

guages such as these have been studied by twentieth-century linguists: they are called pidgins.[18] *Pidgin* is supposed to be based on the Chinese pidgin English word for *business.*[19] When children grow up in this language, a regular grammar is generated and the language becomes a Creole (Pinker 1994, 33). Creoles are rich in vocabulary but poor in inflections. We can look at such languages as a way of turning high-dimensionality optimization problems into a sequence of low-dimensionality optimization problems. In a dynamic programming context, Richard Bellman (1957, ix) called such transformations attempts to avoid the "curse of dimensionality."

### Language Death

As transportation costs fall, of course, there will be more cross-language trading. In the index to their popular book, *Story of English,* Robert McCrum et al. (1986) have more than two-dozen page references to English pidgin to make their case that people are increasingly trading in English-influenced pidgins. What happens to the native language when economic activity moves elsewhere?

Linguists report that somewhere between 10 and 50 percent of the world's languages are dying. One rational choice element in language choice is obvious; as the costs of moving labor falls, parents discover that their children can obtain higher wages in a world language than in their own. But there is something nonobvious about the grammatical trajectory in language death: dying languages become pidginized.[20]

Can our account of Smith's model explain this? Trivially, it predicts that when time is withdrawn from language learning the amount of language learned decreases. When parents decide to move their children into a world language, they do so by talking to them in the world language even if they do not have native competence in it.[21] The children's competence in their parents' language falls; words drop out, and the grammar shifts. To keep with the definition of a language community populated with homogeneous speakers,

---

18. Lingua franca is a pidgin that Smith might have known about (Whinnom 1977).

19. "Many scholars believe that the word *pidgin* was first used for Chinese pidgin English (in which *pidgin* is the word for "business") and was later generalized to mean any language of this type" (DeCamp 1977, 6). According to Robert Hall (1966, xiv): "[O]nly a few hours' trading is necessary for the establishment of a rudimentary pidgin, and a few months or years suffice for the pidgin to assume settled form."

20. Peter Mühlhäusler (1986, 89) observes that "The grammar of a dying language . . . can in many ways be regarded as the mirror image of the grammatical enrichment processes occurring in creolization or pidgin development in expanded pidgins." Lyle Campbell and Martha Muntzel (1989, 191) agree that "Language death may be accompanied by some degree of morphological reduction. . . . While we have several examples in our data, since this is reasonably well established . . . we present only two examples here."

21. Ekkehart Malotki (1983, 622) reports that "For some reason the parents of these children, although perfectly versed in their vernacular, prefer to communicate with their children in English."

students of language death have coined the terms *semispeakers* and *imperfect speakers* to save the phenomenon from being ruled out of court by definition.[22] There is a regular age-competence profile that characterizes dying languages: the oldest speakers—the ones with the most time in the language—speak the most fluently.[23] For instance, the unique mother-in-law language that Dixon documented in Australian languages was only recovered in the memories of the oldest speakers (1989, 144–45, 168). Everyone else had forgotten it.

## Conclusion

Smith's ideas about language choice are consistent with twentieth-century research on pidgin languages. As has been demonstrated, his ideas can be immediately developed to explain the grammatical trajectory of languages as they die, showing the relationship between a pidgin language and that spoken by semispeakers. How naturally and easily these results follow from rational choice considerations argues for taking seriously a wide-ranging rational choice linguistics.

Interesting questions will present themselves in this endeavor. Can we estimate the wage premium elasticity of moving children to a world language? Does the linguistic distance—French is closer to English than Navajo is— matter in the decision to move children from one language to another? Can we make operational the idea that languages are like currency areas?[24] Can one successfully rent seek by impairing movement from one language to another? Can we explain the role of "language mavins"? If every person within a language community is equally fluent, then this activity seems to make no sense (Pinker 1994). If fluency varies with income and education, and income varies with fluency, is there not a possible economic explanation?

---

22. As Nancy Dorian (1981, 115) puts it, "As the language dies, a group of imperfect speakers characteristically appears who have not had sufficiently intensive exposure to the home language, or who have been much more intensively exposed to some other language; and if they continue to use the home language at all, they use it in a form which is markedly different from the fluent-speaker norm."

23. Malotki (1983, 616) found this to be true in his study of the Hopi perception of time: "The present study of Hopi time was accomplished with knowledgeable consultants from an age bracket of approximately forty years and up. Great portions of it could not have been accomplished, however, with informants that are now between twenty and thirty years old. While members from this age group may theoretically still be classified as fluent speakers, they have lost the vital umbilical connection to their linguistic heritage, in particular the traditional knowledge lodged in older Hopi."

24. This analogy was suggested to me by Thomas Borcherding.