Lingua 77 (1989) 187–222. North-Holland

BIOLOGICAL EVOLUTION OF THE SAUSSUREAN SIGN AS A COMPONENT OF THE LANGUAGE ACQUISITION DEVICE

James R. HURFORD

Department of Linguistics, University of Edinburgh, Scotland

Most linguistic theories assume lexicon entries which incorporate the idea of the Saussurean sign, a bidirectional mapping between a phonological form and some representation of a concept. This Sign, like grammars generally, is unbiased with respect to perception or production, and provides part of the cognitive map which speakers use both in speaking and in interpreting the utterances of others. This bidirectionality of the Sign, or code, is a design feature of human language although workable communication does not necessarily incorporate such a feature.

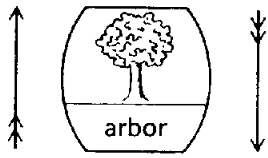
Part of the Language Acquisition Device is a mechanism which mentally constructs such a bidirectional mapping, on the basis of observed samples of communicative behaviour (transmission and reception). This basic aspect of the LAD presumably evolved because of its superiority over other conceivable mechanisms for acquiring a basis for communicative behaviour from a sampling of observed transmission and reception data.

Three conceivable strategies for acquiring the basis of communicative behaviour, here labelled the Imitator, Calculator, and Saussurean strategies, are defined as functions from samplings of observed behaviour to acquired behaviour patterns. Essentially, the Imitator imitates the transmission and reception patterns in the observed sample; the Calculator constructs optimal reception responses to the transmissions in the observed sample and optimal transmission responses to receptions in the observed sample; the Saussurean imitates the transmission behaviour in the sample, but shapes his reception behaviour to mirror the acquired transmission behaviour, thus internalizing the equivalent of a set of bidirectional Saussurean signs. Extensive simulations of populations endowed with these innate strategies show the Saussurean strategy to be a winner in evolutionary terms; it produces individuals who communicate more successfully than individuals endowed with the other two strategies.

1. What is to be explained

I assume that the fundamental formal structure underlying human languages is the bi-directional Saussurean sign, as in (1), representing a relation between a concept and the sound-image of a word.

(1)



(Saussure (1959: 67))

The two arrows in Saussure's diagram indicate that the sign is neutral as between speaker and hearer. The sign is an independent piece of information, mentally represented, I presume, available for use in both transmission and reception.

Of course, a developed human language is not *merely* a finite set of Saussurean signs mapping simple concepts to simple expressions. A developed language has a recursive syntax and associated semantic rules, allowing formation of an infinite set of complex signs, mapping complex concepts (propositions) to complex expression-images (phonological representations of sentences). In addition a developed language has a full apparatus of rules at other levels, e.g. phonological rules, morphological rules, etc. The essential and basic relation in the complex lexical entries of generative grammars of developed languages is the Saussurean *signifié-signifiant* relation. Presumably the phylogenetic basis for modern developed complex language was a simple one-word stage, without syntax, but *with* the Saussurean two-way relationship between a concept and a word-image, a *signifié* and a *signifiant*. In this paper I describe a model of the evolution of this early stage of language, a model which explains it.

What could *explain* the fact that the Saussurean sign is the key formal structure at the foundation of languages? The fact itself is so obviously self-evident that any notion of explaining it might at first seem bizarre or doomed to triviality. But such facts as that heavy objects fall to earth and that animals and plants inherit the characteristics of their parents were also once held to be inscrutably axiomatic or downright unexplainable. We now have explanations for these facts (gravity/relativity, genes/DNA) and see that they were not so axiomatic or inescapable after all: these things might conceivably have been otherwise, except that we can now see that they follow as consequences of rich theories with the power to explain, and even predict in detail, much else besides. I do not claim to present a theory of the depth of Newton's or Mendel's, but their example can perhaps open the door to the idea that some antecedent combination of circumstances determines the central role of the Saussurean sign in language.

The explanation to be presented has a nativist component, in the spirit of Chomsky's proposals for explaining the formal organization of languages. Such nativist explanations account for aspects of the form of languages by appealing to characteristics of the innate Language Acquisition Device (LAD) in each newborn child. Languages are the way they are, this style of explanation goes, because each possessor of a language initially takes hold of it in the only way he knows how. Children are biologically programmed to represent the acquired knowledge of their language in specifically determined ways, and thus we find that languages known by adults are objects with just such specifically determined formal characteristics. The form of the system known, i.e. the structure of the language, is only partially (for Chomsky quite trivially) determined by structure apparent in the input experiences of the acquirer, and partially (most significantly for Chomsky) determined by innate predispositions to acquire knowledge structured in a certain way.

That languages are structured according to the bi-directional Saussurean sign is not apparent from direct observation of the superficial data of adult performance. The acts a child observes are acts of transmission and (other) acts of reception. Allow for the moment that a child somehow has access both to a spoken expression and to its intended meaning in the mind of its utterer. On witnessing an act of transmission by an utterer, the child is not logically required to assume that this same adult would relate the same expression to the same meaning in an act of reception. To infer patterns of reception from observed patterns of transmission, or vice versa, is to go beyond what is strictly deducible. Yet I have no doubt that the child language acquirer naturally comes to just this conclusion, and to no other. The language acquirer knows that his business is to internalize a system of Saussurean signs, neutral as between speaker and hearer, bi-directionally relating *signifié* with *signifiant*, and he induces this system from suggestive observations from which it does not logically follow.

It is worth amplifying the sense in which observation (sampling) of transmission and reception can be separate and independent. For the most extreme, but least typical case, imagine observing a creature alone, beyond signalling range of any of its fellows. Observation of transmission here could consist in seeing the creature concentrating on some object, and then emitting some signal. Conversely, observation of reception could consist in observing the creature perceive some signal, and then focussing its attention on some object. Less drastically isolated situations can be envisaged, in which fellow creatures are not necessarily absent, but where distinct events of transmission and reception can still be isolated. For instance, a young creature might observe its mother react to a wasp with a particular excited noise, while the rest of the group continues nonchalantly grooming, ruminating or sun-bathing, paying no attention to the wasp or the mother; this would be observation of isolated transmission. Crucially here it is assumed that the youngster becomes aware of the wasp *not as a result* of the mother's noise, but independently of it (ideally before it) yet is able to conclude that somehow the presence of the wasp brought about his mother's action. For creatures with any general appreciation of cause and effect, however unconscious, this is not at all implausible. Or the young creature might itself accidentally produce a certain noise, and find that its mother immediately acts as if reacting to some specific kind of object (say a wasp); this would be observation of isolated reception.

Presumably, a young creature born into a group with developed successful adult communication would typically observe many encounters in which transmission and reception were not isolated, but coordinated. An example would be a youngster seeing a wasp, noting his mother react to it with a particular signal, and noting the other adults reacting to the signal by attending to the wasp. It would still be possible in observing such cases for the observations of transmission and reception to be independent, although based on the same encounter taken as a whole. The important thing would be for the learner to be able somehow to identify the object triggering the transmission of the signal *as* triggering the signal, and the reactions to the signal *as* reactions to the signal.

Being competent human language-knowers, our problem as researchers (or at least speculative theorists) into the origins of signalling systems is to imagine what it is like *not* to know an inherent connection between transmission and reception. Two of the learning strategies to be compared later in this paper have no intrinsic connection between transmission and reception built into them, while a third does have, in the sense that sampling of transmission alone is taken as a sufficient basis for the acquisition of coordinated transmission and reception.

Chomsky's nativist paradigm for explaining aspects of language form is classically represented as in (2).

(2)

Primary linguistic data	→	Language Acquisition Device	→	Individual grammatical competence

In this, the form of a speaker's acquired competence is the *explanandum* and the LAD is posited as the *explanans*. To many, merely saying 'the answer lies in the Mind' *tout court* has not seemed sufficient explanation, particularly because we have no independent access to formal characteristics of the Mind. Further, we are offered no plausible reason why the Mind should have come to be structured in the specific ways suggested by Chomskyan linguists. Functional explanations for the alleged innate structure of the LAD are typically not proposed. In this paper, the nativist mechanism is only one step in the explanation of the Saussurean sign. An evolutionary functional answer is offered for the deeper question of what explains the *explanans*, the fact that the Saussurean sign is such a basic ingredient of the LAD. The form of answer that will be given is represented in (3), where the LAD appears as the *explanandum*, itself the output of antecedent mechanisms.

(3)

Biological Mutations → | FACTORS INVOLVED IN SUCCESSFUL COMMUNICATION IN THE HUMAN ENVIRONMENT | → Language Acquisition Device

What were the evolutionary mechanisms by which humans came to be capable of, and predisposed to, mentally representing (and using for communicative purposes) the particular formal relationship between a concept and a word-image that is encapsulated in the Saussurean sign? I assume that, prior to this evolutionary step, the ancestors of Man were complicated enough to license us to talk about concepts or mentally represented meanings, whose relation to overtly observable behaviour may be quite indirect. I assume that such 'pre-Saussurean' speakers and hearers possessed stores of concepts and that their attention may or may not at any given time have been focussed on particular concepts. Concepts would be of recurring types of objects and events in their environment, including the observable actions of their fellows. I assume it is possible to give an account (perhaps ultimately in behavioural terms) of what it is for a concept possessed by one individual to be the same as a concept in the mind of another. Building on these assumptions, I define as an instance of successful communication any encounter between individuals where one (the transmitter), while mentally attending to a particular concept, carries out some observable act (which may be a gesture, a vocalization, or whatever), and another individual (the receiver), as a result of observing this act comes to attend to the same concept.

The processes of transmission and reception are logically distinct and *not necessarily* (or 'logically') coordinated with each other. There is no logical necessity for a creature attending to some concept, and wishing to convey it to another, to make the very same gesture which, if made by another, would convey that concept to him. The question posed here is how, as a *contingent* matter, this coordination could come about. This, surely, is how the problem must be set up in a discussion of the origins of communication by symbolic behaviour. The typical case of successful communication that I am concerned with involves coordination of transmission and reception within an individual in such a way that the individual could communicate perfectly with others exactly like him. But it will illuminate the logical relationship between transmission, reception, and communication to look at a case where transmission and reception can be well coordinated allowing successful communication of a limited sort to take place even though a given individual could not communicate with another exactly like him. With separate transmission and reception devices, the following situation is possible.

(4) Successful non-Saussurean communicative scenario

	Transmission: concept to gesture mappings			*Reception*: gesture to concept mappings		
Individual A	Concept	*a*	*b*	Gesture	1	2
	Gesture	1	2	Concept	*b*	*a*
Individual B	Concept	*a*	*b*	Gesture	1	2
	Gesture	2	1	Concept	*a*	*b*

Here, if individual A is attending to concept *a* and wishes to convey it to individual B, he emits gesture 1, and on reception of this gesture, individual B brings concept *a* to mind, as required; and similarly for either concept communicated by either individual. That is, in this hypothetical and artificially limited scenario, individuals A and B can communicate perfectly with each other, but clearly the concept-to-gesture and gesture-to-concept mappings that each has internally represented do not amount to a knowledge of a two-way Saussurean sign.

The assumption that 'a language is a code which pairs phonetic and semantic representations of sentences' (Sperber and Wilson (1986: 9)) or, in other words, that the basis of human language *is* the two-way formal relationship encapsulated in the Saussurean sign, portrays a language as a kind of formal object, represented in the minds of speakers, and drawn on by them in active use. This view distinguishes the formal object, the language system or code, from its use. Such a distinction between language system and language behaviour is standard, emerging in slightly different guises in the Saussurean *langue/parole* dichotomy and the Chomskyan *competence/performance* opposition. I wish to emphasize the difference between knowledge of a linguistic system and a capacity for calculating intelligent responses to events in one's environment, including the communication-related behaviour of fellow members of one's social group. The point is to distinguish between a general capacity to act intelligently and language structure (whose use may well involve considerable intelligence). To be capable of possessing human language is not just to have reached a particular high level of intelligence. It is something more specific, and more structured.

Consider again the 'Successful Non-Saussurean Communicative Scenario' presented above. Here the concept-to-gesture and gesture-to-concept mappings known to each individual could well be seen as intelligently calculated to yield optimal responses to the behaviour of the other individual, resulting in successful communication. That is, such a successful situation could have come about *because* individual A expressly tailors his transmission behaviour to coordinate with what he observes to be individual B's reception behaviour, and vice versa. Here, each individual would be second-guessing the responses of the other, and devising a scheme for his own behaviour calculated to produce maximally successful communication between them. The chicken-and-egg problem of how such a process could get started might be solved by allowing that the initial behaviours were random, and that the coordinated final behaviours emerge from a succession of acts intelligently calculated to yield optimal communication. But these separate internalized schemes for transmission and reception do not add up to internalization of a two-way Saussurean sign. Each individual's concept-gesture mapping is different for transmission from what it is for reception. Yet these individuals communicate perfectly. They manage it in a non-Saussurean way. The hypothetical Successful Non-Saussurean Communicative Scenario (4) involves just two individuals, two concepts, and two gestures. But could such success, based only on individuals intelligently calculating optimal responses to each others' behaviour, be generalized to more complex situations, involving many individuals, many concepts and many gestures?

2. Sociobiological explanation

The basic structure of the argument to be presented is as follows:

(*1*) Individuals who are more successful communicators enjoy a selective advantage and are more likely to reproduce than individuals who are worse communicators.

(*2*) An innate Saussurean strategy for acquiring a communication system is superior to other conceivable strategies, and its possessors tend to enjoy a reproductive advantage over others, thereby increasing the prevalence of innate Saussurean strategists in the next generation.

(*3*) Therefore, over an evolutionary timespan, the Saussurean strategy displaces all rivals, and ends up being *the* strategy by which communication systems are naturally acquired.

In this section I introduce this type of 'sociobiological' explanation and discuss its applicability to the linguistic case.

There are so many interacting elements in the structures and environments of living organisms that it is notoriously impossible to *prove* that possession of some behavioural characteristic necessarily confers advantage. But this fact need not inhibit the production of plausible speculative theories, which can be enlightening and yield as much understanding as one can hope for in the area of evolution. For guidance in appropriate model-building I turn to a grouping of ideas, known as 'sociobiology', that has attracted much attention in the field of animal behaviour over the past decade and a half (see Clutton-Brock and Harvey (1978), KCSG (1982) for representative samples of introductory and more advanced work). For linguists, some of whom may be only marginally aware of this work, I sketch below the main aspects of sociobiological work relevant to possible explanations of linguistic structure.

In some circles, sociobiology has become controversial, and the enterprise itself has been attacked in critiques sometimes tinged, sometimes drenched, with ideology. Three of the more responsible and less hysterical critiques are Kitcher (1985), Rose, Lewontin and Kamin (1984), and Reynolds (1980), although even these, it seems to me, tend to attack straw-man conclusions which most practising sociobiologists do not, and would not, draw from their own work. This is not the place to join the sociological/philosophical/ethical debate on whether sociobiological models can be applied to human social behaviour. Obviously, I believe that they can, and I ask the reader to consider the present proposal relating to language on its own merits, and without reading into it more than is actually being claimed. I do not believe that the contribution of culture and tradition to human nature is negligible. In fact, I have argued elsewhere (Hurford (1987)) for a view of language which incorporates *both* the biological/psychological and the social/cultural/historical determinants of language form. The interesting research questions lie in investigating the *interactions* between biology and culture. This work is intended as a contribution in that spirit.

In essence, sociobiology, though sometimes hailed as revolutionary, is solidly traditional in its adherence to Darwinian evolutionary theory and classical Mendelian genetics. Whereas classical work has tended to concentrate on the physical characteristics (shape, size, colouring, organs, etc.) of individual creatures, such as the mammalian eye and the wings of the fruit fly, sociobiology investigates traits in the behaviour of individuals in relation to other members of their social group. Prominent examples of such behavioural traits include the instinct to keep spatially close to one's fellows (the 'herding instinct', cf. Hamilton (1971)), aggression or docility in potentially competitive encounters with one's fellows ('hawk' and 'dove' strategies, cf. Maynard Smith (1976)), the giving of alarm calls (cf. Maynard Smith (1965)), mating behaviour (cf. Orians (1969)). The novelty of sociobiology is in the construction of precise mathematical models of the interaction of individuals with specific innate dispositions to social behaviour.

A key notion in sociobiology is that of *a strategy*. Strategies belong to individuals and are precisely defined propensities to act in certain ways in relation to other members of the social group. Strategies may be very simple (e.g. 'always fight') or more complex, perhaps conditional on judgements about the relevant situation (e.g. assessments of a rival's strength and willingness to fight). In practice, sociobiologists adopt the same attitude to the strategies they discuss as Chomskyan linguists do to the hypothesized properties of the innate Language Acquisition Device. That is, they undertake no responsibility to provide physiological or neural correlates, although they do not deny that there must be such correlates, since both the strategies and the properties of the LAD are genetically transmitted.

Linguists and sociobiologists work in domains whose entities are at a level of abstraction higher than physical neurons and synapses. The linguist manipulates theoretical entities such as rules (phrase structure rules, phonotactic rules, semantic rules of inference, etc.), and categories (e.g. Noun, Verb, tense, number, gender). These theoretical entities belong to systems which, taken in their entirety, are reflected in actual behaviour, but the precise physical causal mechanisms are not spelled out. It is assumed that there is value, in the present state of ignorance about neurological processes, and especially about the genetics of such processes, in working at the abstract level. Similarly, when a sociobiologist hypothesizes some genetically transmitted strategy, he feels under no obligation to spell out how this strategy is represented in the DNA. A careful sociobiologist typically avoids identifying particular hypothesized strategies with single genes. For example, he does not speak of a 'gene for fighting or fleeing according to the apparent strength of one's rival'. Instead, he hypothesizes simply that individuals can somehow be endowed by their heredity with dispositions to react in specified ways to specified situations. As there are no grounds for any particular more complex assumption about the genetic transmission of behavioural traits, it is assumed that such dispositions can be passed on to offspring undistorted. The assumption that strategies correspond to sets of genes is, however, often the only one practicably available in constructing idealized mathematical models and simulations. Such idealized mathematical modelling and simulation are pursued for the insights they provide into the simple principles underlying the undoubtedly more complex interactions of genetics and behaviour. In this sense, both sociobiology and generative grammar, interpreted as the study of the Language Acquisition Device, operate at the same high level of abstraction.

The normal everyday sense of the term 'strategy' is well adapted to many of the situations, such as conflict over a resource, behaviour in the presence of predators, mating, etc., discussed by sociobiologists. This is perhaps because the term 'strategy' in its everyday sense carries an implication of observable manifestation; an observer can, in some sense, see what strategy a subject is following. Clearly, to match theory with observation in the case of non-humans, definitions of hypothesized strategies must relate to observable behaviour. The predicates, such as 'fight', 'flee', 'give alarm call', 'adopt submissive position', etc., used in connection with strategies discussed for non-humans must have at least roughly agreed-upon real-world denotations, or the theory simply would not relate to the animals it purports to be about. And perhaps, for some such strategies, it might even be sufficient to assume no more than a simplistic behaviourist stimulus–response mechanism in the animal.

But the case of human language acquisition is different. It would be a perverse use of the terms to say that the primary linguistic data to which a child is exposed is a 'stimulus', and the child's 'response' is to acquire the language. Stimulus-response theories of behaviour typically posit no internal representation, or knowledge, on the part of the behaver. Some sociobiological models also, though they do not arise from behaviourist assumptions, posit no internal representation or knowledge in the behaver. But the more interesting and plausible strategies discussed by sociobiologists require the capacity to retain knowledge, for example the strategy of reciprocal altruism discussed by Trivers (1971), which requires that a creature be able to remember the past behaviour of other individuals towards it. In the case of language-acquisition, the formation of internal representations becomes the central issue. Language-acquisition is the acquisition of a complex set of internal representations, or (subconscious) knowledge of the grammar of a language. The manifestation of this knowledge in observable behaviour is quite indirect. So the use of the term 'strategy' in the context of language acquisition requires some discussion.

I shall keep the term 'strategy' and use it as follows. Where the Language Acquisition Device is a complex function from primary linguistic data to internalized knowledge of the grammar of a language, a strategy is a component, a subfunction, of this function. Each strategy or subfunction yields a particular component of the individual's total knowledge of his language. In the complex structure of modern evolved language, for instance, there is phonological knowledge of the phonological units (phonemes, say) and phonological rules of a language, syntactic knowledge of the rules determining word-order, semantic knowledge of the functions determining the meanings of sentences, etc. A language acquisition strategy is a disposition to internalize representations of a particular form. Individuals possessing different strategies, but exposed to the same primary data, may internalize different representations, with different (indirectly) empirically available consequences. The programme of research followed in this paper starts with primitive protolinguistic cases, where extremely simple structure is attributed to members of the evolving species, for instance without the double articulation into phonology and syntax characteristic of modern evolved language, or even any kind of recognizable syntax or phonology at all. The aim is to demonstrate that specific kinds of internal representation, as opposed to other conceivable kinds, give rise to greater communicative possibilities, and hence to evolutionary advantage. Once a species with a particular innate structuring of internalized knowledge relative to the communication systems it can acquire has become established, other successively more complex possibilities can arise. The research plan is to advance beyond arguments involving the evolutionary advantage of specific very simple structures, such as the basic Saussurean sign, towards arguments involving the evolutionary advantage of more complex structuring resembling the modern evolved case. Work in a closely related vein on the biological evolution of the critical period for language acquisition is reported in Hurford (to appear). The present paper reports results relating to the disposition to internalize knowledge about one's communication system in the format of the bi-directional Saussurean sign.

The use of the word 'strategy' here does not, of course, imply that a strategy's possessor has any conscious or intentional control over the relevant actions, but equally there is no implication that reasoned action is not involved in some sense. The lack of involvement with animals' possible reasons for their actions makes the sociobiological approach seem, to some, particularly inappropriate to *human* behaviour. But this cannot be taken as an objection to a sociobiological approach to the properties of the innate LAD, because intentionality, reasoned action, and conscious control are not involved *in the relevant sense* in the task of language acquisition. Clearly, a child has reasons for producing particular utterances; she has a degree, perhaps a high degree, of intentional control over the specific utterances she produces in specific situations, but only in the *meaningful* aspects of the utterance. For instance, the child will choose which words to use, because they (already) mean what they do in the language she is acquiring, but she cannot choose or in any significant way determine what the individual words mean. (So Alice was mainly right, and Humpty Dumpty was mainly wrong.) Similarly in grammar, the language acquirer can choose an appropriate construction for use on a particular occasion from the set of constructions made available by the language, but typically has no say in specifying this set; that has already been determined for her by centuries of the spiral through the twin filters of the LAD and the social arena of language use. A child does not have conscious control over the innate structure into which she fits the language she acquires. To a relatively minor extent, an individual can invent novel variations in the language she acquires, but these inventions can never fall outside the set of possibilities determined by the individual's innate apparatus for representing linguistic knowledge; and it is with the innate structure that the proposed sociobiology of language is concerned.

I shall assume that the creatures to be discussed are capable of reasoned action directed towards practical goals that they have, somehow, conceived and internally represented. The courses of action they adopt in the pursuit of those goals may include communicative action with their fellows. As an initial idealization, I shall assume that all creatures in a population have identical capacities for reasoning from goals to appropriate communicative action. That is, there will be no difference between individuals as to which message they choose to transmit in order to achieve a given goal. Likewise, there will be no difference between individuals as to the appropriate action in response to a particular message received. And, barring mutations, there will be no differences between individuals in the range of goals they can conceive of, nor in the range of messages they can transmit and receive.

Given the basic notion of a strategy, a further key notion in sociobiology is that of the 'invasion' of a population by a strategy. Say a mutant with a novel strategy is born into a population whose other members lack this strategy. Since most mutations confer no advantage on an individual, in the typical case the mutant will die without issue. But sometimes a mutation is advantageous, allowing the mutant to survive better in the environment than its non-mutant fellows. Assuming the mutant can breed with its non-mutant fellows, it is likely that the mutant gene(s) will be passed on to subsequent generations and, over the generations, become more widespread, possibly ousting the non-mutant allele(s) entirely. Where this happens, one deems the mutant gene(s) to have successfully invaded the population.

An example of an invasive strategy is the herding instinct discussed by Hamilton (1971). Given a group of animals preyed on by predators which appear at random and strike at the nearest prey animal (an obvious idealization), each prey animal has a 'domain of danger', defined as the space containing all points nearer to it than to any of its neighbours. An animal with a large domain of danger has a proportionately lower chance of surviving. A strategy minimizing the domain of danger will enhance an animal's chances of survival. One such strategy is to move to a point where one's average distance from one's fellows is as small as possible. Animals which do this will tend to survive better than animals which don't, and this strategy will invade the population.

If all animals repeatedly move to points closer to their fellows, they will eventually congregate in closely packed herds. At this point, the idealization breaks down, because animals need a certain amount of free space to forage for food. Now the modelling of behaviour needs to become more complex, taking into account other considerations besides the avoidance of predators, factors such as the availability of food, water, air, etc. But the simple model has shown us something about the herding instinct and why it evolved. Evolution is an arms race, as Dawkins (1986) has put it, with one strategy prevailing until it is superseded by an even more successful one, or until the species' enemies evolve successful counter-strategies. The simple species of predators who appear at random and strike at the nearest pray animal could 'develop' (unconsciously, over an evolutionary timespan, of course) a successful counterstrategy of looking for herds of prey animals and striking at their middles. But this new predatory strategy depends on there being herds in the first place.

Another example of an invasive strategy is that of 'hawk' in a population of 'doves' (Maynard Smith (1976)). 'Hawk' is defined as the strategy of attacking a rival for a resource, and trying to fight him off. 'Dove' is defined as the strategy of patiently trying to outwait a rival for a resource. In an encounter between two 'doves', each will spend a long time simply eyeballing the other, waiting for it to give up the resource. In an encounter between two 'hawks', there will be a fight, at some cost to both, but with an unequal distribution of benefit. In an idealized model, 'hawks' can expect to win half of their encounters with other 'hawks', at a certain average cost and with a certain average benefit per encounter. Similarly, 'doves' can expect to win half of their encounters with other 'doves', again with some average cost and benefit per encounter. In a 'hawk'-'dove' encounter, however, the 'hawk' will always win the resource, and reap benefit at less cost than in an average 'hawk'-'hawk' encounter; and the 'dove' will always lose, incurring cost with no compensating benefit. If the cost to a 'dove' in a 'hawk'-'dove' encounter is greater than the average benefit-minus-cost in a 'dove'-'dove' encounter, say because the 'hawk' always attacks its opponent regardless of whether it puts up a fight, then a population of 'doves' would be rapidly invaded by a 'hawk' mutation. Depending upon precisely how one defines the cost incurred in 'hawk' and 'dove' encounters, however, it is not inevitable that a broadly hawkish strategy will invade a population of 'doves'. Different mathematical details will yield different balances between strategies.

The models set up by sociobiologists typically involve comparison between different hypothesized strategies, e.g. 'hawk' versus 'dove'. The technique gives an insight into the advantages of a specified strategy or strategies in the presence of other strategies. Change the standard of comparison or the assumed background of other strategies, and one has a different model, suggesting other (but not, of course, contrary) conclusions. One can conclude that strategy S1 will always successfully invade a population characterized by strategy S2, but the introduction of a third strategy S3 into the scenario may change the picture entirely, with S2 and S3 jointly ousting S1. In choosing scenarios of strategies to theorize about, the sociobiologist has a responsibility to maintain as much verisimilitude as possible, despite the obvious methodological need for idealization. Otherwise the conclusions drawn from the idealized models would provide no insight into the real-world situations they purport to model.

In a direct application of the sociobiological technique to the case of the Saussurean sign, I will define a simple model of communication between members of a group. Within this framework, I define three innate strategies by which newborn individuals might acquire from adult performance the knowledge required to participate in the communicative activity of the group. These strategies are labelled the Imitator strategy, the Calculator strategy, and the Saussurean strategy. As the reader will have guessed, the conclusion will be that the Saussurean strategy proves superior to the others. But before this dénouement can be acted out in full view, the stagehands must use the next section as an entre-acte to erect a visible, plausible, and computationally tractable characterization of the essence of communicative behaviour.

3. Communicative potential and interpretive potential

The formal exercises in this section provide a rigourous way of characterizing the transmission and reception behaviour of individuals and populations and define the degree of success with which they can communicate with each other. A grasp of the basic notion that the probabilities of events can be expressed as numbers between 1.0 (for absolute certainty of occurrence) and 0 (for absolute certainty of non-occurrence) will probably be enough for the mathematically fearsome to get through the section; for others, it should be plain sailing.

For any pair of individuals, s and h, an index of s's 'communicative potential' relative to h and relative to a given (concept of an) object o can be expressed as $\mathrm{CP}(s,h,o)$ and defined as follows.

(5) $\mathrm{CP}(s,h,o)$

$$=_{\mathrm{def}} \sum_{i=1}^{i=n} [\mathrm{P}(s \text{ transmits signal } i \text{ for } o) \times \mathrm{P}(h \text{ interprets } i \text{ as object } o)],$$

where n is the number of possible signals 'in the system'. Thus, $\mathrm{CP}(s,h,o)$ is the probability, given that s is attending to o, of s 'successfully referring' to o in an encounter with h. The converse of CP is IP (interpretive potential) $\mathrm{IP}(h,s,o) = \mathrm{CP}(s,h,o)$. $\mathrm{IP}(h,s,o)$ is the probability, given that s is attending to o and transmitting a 'signal for' o, of h interpreting that signal as corresponding to o.

A more general measure, CP(s,h) can be defined, not relative to any particular object, but averaging over all objects 'in the system'.

(6)

$$\mathrm{CP}(s,h) =_{\mathrm{def}} \frac{\sum_{i=1}^{i=n} \mathrm{CP}(s,h,i)}{n},$$

where n is the number of objects in the system. CP(s,h) is the probability, given that s is attending to any arbitrary object, of s successfully referring to that object in an encounter with h. Conversely, IP (h,s) = CP(s,h). IP(h,s) is the probability of h successfully interpreting any 'signal' emitted by s. These definitions of 2-place CP and IP assume that all objects are in some sense equally important or equally frequent subjects of communication; if this assumption is deemed unwarranted, alternative, more complex definitions of 2-place CP and IP can be formulated.

(The discussion assumes that an answer can be given to the question of how many objects are 'in the system' for some group of potentially communicative creatures. As a first approximation, this may be thought of as the number of objects/situations for which the creatures have (some degree of) categorial perception.)

Useful measures are an individual's CP and IP relative to the whole population. CP(s) is the probability that an utterance by s will be successfully interpreted by any arbitrary individual from the whole population. IP(h) is the probability that h will successfully interpret any arbitrary utterance by an arbitrary individual from the whole population. An individual's CP relative to the whole population may be obtained by averaging over his CP relative to all other individuals in the population. This assumes that all other individuals are equally frequent interlocutors; again, if this assumption is rejected, more complex alternative definitions of 1-place CP and IP can be formulated. An individual's communicative and interpretive potentials, relative to the whole population (or indeed relative to any subgroup or other single individual) are not necessarily the same – see below.

An even more general measure, relating to the whole population, is CP. CP is the mean of all values of CP(s,h) for all speaker-hearer pairs $\langle s,h \rangle$. CP is the probability that an arbitrary encounter between arbitrary individuals in the population will involve successful reference. In general, where the measure is of the whole population, CP = IP.

In general, for two given individuals, X and Y, CP(X,Y) = IP(Y,X). But note that, given separate mechanisms for 'communicating' (transmission) and 'interpreting' (reception), an individual may communicate completely successfully, but interpret with only a chance level of success. An example follows.

(7)

	Transmission: object to signal mappings						*Reception*: signal to object mappings					
	Object	*a*	*b*	*c*	*d*	*e*	Signal	*v*	*w*	*x*	*y*	*z*
Individual X		$\vert$	$\vert$	$\vert$	$\vert$	$\vert$		$\vert$	$\vert$	$\vert$	$\vert$	$\vert$
	Signal	*v*	*w*	*x*	*y*	*z*	Object	*a*	*b*	*c*	*d*	*e*
	Object	*b*	*c*	*d*	*e*	*a*	Signal	*v*	*w*	*x*	*y*	*z*
Individual Y		$\vert$	$\vert$	$\vert$	$\vert$	$\vert$		$\vert$	$\vert$	$\vert$	$\vert$	$\vert$
	Signal	*v*	*w*	*x*	*y*	*z*	Object	*a*	*b*	*c*	*d*	*e*

CP(X,Y) = IP(Y,X) = 1.0 CP(Y,X) = IP(X,Y) = 0.0

In this situation, individual X will always communicate successfully to individual Y, and (by definition) Y will always successfully interpret X's utterances. But Y will never communicate successfully to X, and, correspondingly of course, X will never successfully interpret Y's utterances. Other, intermediate situations exist, where the probabilities of emitting and interpreting signals fall between 1.0 and 0.0. In the great majority of logically possible situations, for arbitrary X and Y, CP(X,Y) $\neq$ IP(X,Y).

Given separate mechanisms for transmission and reception, communicating and interpreting, an individual would not necessarily be able successfully to communicate with or interpret his identically behaving twin, though both have identical responses to objects and signals. But, for the special case of identically behaving twins, X and Y, CP(X,Y) = IP(X,Y), whatever (high or low) value CP(X,Y) and IP(X,Y) may happen to have.

In order to investigate the theoretical properties of CP and IP, as well as to conceptualize them more clearly, one can draw up matrices of transmission and reception probabilities for the individuals or groups involved. Here are some hypothetical probabilistic examples.

(8) Individual (or group) X transmission

		Signals transmitted → w	x	y	z
Objects causing transmission ↓	a	0.2	0.3	0.4	0.1
	b	0.4	0.3	0.2	0.1
	c	0.5	0.5	0	0
	d	0.25	0.25	0.25	0.25
	e	0.333	0.333	0.333	0

Here the number in entry (a,w), for example, is the probability of this individual giving signal w while attending to object a. The figures in each row sum to 1.0. The columns do not necessarily sum to 1.0.

(9) Individual (or group) Y reception

		Objects associated with signal by receiver → a	b	c	d	e
Signals received ↓	w	0.7	0.1	0.05	0.15	0
	x	0.1	0.5	0.1	0.2	0.1
	y	0.15	0.1	0.2	0.25	0.3
	z	0	0.2	0.3	0.3	0.2

Here the number in entry (w,a), for example, is the probability of this individual having his attention drawn to object a when another individual gives signal w. Again, the figures in each row sum to 1.0; the columns do not necessarily sum to 1.0.

Given two individuals (or groups) X and Y, CP(X,Y) can be calculated from such matrices in the following way. Let X's transmission matrix be T, and Y's reception matrix R. CP(X,Y) is the mean of the entries in the main diagonal of the product TR of the two matrices (as defined in linear algebra). For example, the product of the two matrices above is:

(10)

		Objects associated with signal by receiver →				
		a	*b*	*c*	*d*	*e*
Objects causing transmission ↓	*a*	0.23	0.23	0.15	0.22	0.17
	b	0.34	0.23	0.12	0.2	0.11
	c	0.4	0.3	0.075	0.175	0.05
	d	0.237	0.225	0.162	0.225	0.15
	e	0.316	0.233	0.117	0.2	0.133

I will call such a matrix the 'communication matrix for (X,Y)'. (A communication matrix for (X,Y) is also an 'interpretation matrix for (Y,X)', since CP(X,Y) = IP(Y,X). To obtain such communication and interpretation matrices, the transmission and reception matrices must be multiplied in that order; matrix multiplication is not commutative.) A communication matrix for (X,Y) gives the probability, for any pair of objects, O1 and O2, of X attending to O1 and producing a signal which causes Y to attend to O2. The entries in the main diagonal give the probability of successful reference for each object, i.e. the cases where O1 = O2. In the situation of perfect communication from X to Y, the product matrix is the identity matrix, in which all the entries in the main diagonal are 1, and all other entries are zero. In the situation of purely random relations between objects and signals, the communication matrix for a pair of individuals will be homogeneous, with all entries the same, the inverse of the number of objects 'in the system', e.g. 0.5 for 2 objects, 0.2 for 5 objects, and so on. I will call such a matrix the 'chance communication matrix'.

The 'biases' in transmission and reception matrices are what give rise to the CP for the system as a whole. A row can be biased in as many ways as it has entries; that is, its highest number can be the 1st entry, or the 2nd, or the *n*th. A maximally biased row has the value 1 in one entry and 0 in all others. When the biases in a given pair of transmission and reception matrices are both coordinated and maximized, the overall CP is at its global maximum. The biases of a transmission matrix T are coordinated with those of a reception matrix R when for each row *r* in T the highest number is in column *c*, and in R the highest number in column *r* is in row *c*.

Where there are more objects than potential signals, then of course the overall CP cannot rise to the maximum of 1.0, as not all signals can be shared out unambiguously between the objects; in such a situation, there are 'not enough signals to go round', and the global maximum CP for the whole system remains below 1.

4. Imitators, Calculators, and Saussureans: three learning strategies

Consider newborn creatures who arrive in a world with transmission and reception of signals already happening about them. If they are not innately endowed with any knowledge of object–signal correspondences, they need some strategy for extracting from the behaviour they observe around them guidelines to direct their own future behaviour, if they are to become communicating members of the group. As throughout, I assume that acts of transmission and reception are separate, although, in a successfully intercommunicating group, they will be coordinated. The eventual successful coordination of transmission and reception does not necessarily make a single thing of them, just as scoring a bulls-eye at shooting does not conflate the squeezing of the trigger and the penetration of the target by the bullet. This is not to assume that the successful initiate into a scene with coordinated transmission and reception does not *make* of the coordinated transmission and reception a single higher level entity, a mental, psychological, or perhaps even social, entity, but neither does my approach take such construction of a higher level entity for granted. Indeed, the issue addressed in this paper is exactly that of whether any evolutionary advantage can be plausibly seen in a learning strategy which does involve the internal construction of a Saussurean sign. Such an internalized sign *stipulates* the coordination of transmission and reception in the learner's own behaviour, rather than merely relying on their coordination in the behaviour of the group.

We deal with simple one-word utterance scenarios. Communicative behaviour goes on all the time, and there is no way a learner can collect all of it. He has to take a finite sample, and if his learning strategy provides an adequate function from the sample to his own behaviour, all will be well. In fact, young children seem to acquire new lexical items with extraordinary facility, often on the basis of even a single exposure. The models to be explored here assume that such single exposures to adult behaviour are sufficient to allow the learner to shape his own behaviour decisively. As with everything else here, this is of course an idealization. In the model to be explored, all learners are exposed to one randomly chosen transmission of a signal for each object involved, and, depending on the learning strategy, one randomly chosen interpretation as an object of each signal involved. The random choices are weighted according to the probabilities inherent in the mean transmission and reception behaviours of the whole population.

I will give a concrete example, working with a 'system' containing 5 objects and 7 possible signals. Say the mean population transmission and reception behaviour are as in these illustrative matrices:

(11) *Transmission* *Reception*

$$\begin{bmatrix} 0.1 & 0.2 & 0 & 0 & 0.4 & 0 & 0.3 \\ 0 & 0.1 & 0.1 & 0 & 0.8 & 0 & 0 \\ 0.2 & 0 & 0 & 0.3 & 0.1 & 0.2 & 0.2 \\ 0 & 0 & 0.7 & 0.1 & 0.1 & 0 & 0.1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \qquad \begin{bmatrix} 0.2 & 0.2 & 0.2 & 0.2 & 0.2 \\ 0.1 & 0.2 & 0.3 & 0.4 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0.1 & 0.9 & 0 & 0 & 0 \\ 0 & 0 & 0.3 & 0.3 & 0.4 \\ 0.8 & 0 & 0 & 0.1 & 0.1 \\ 0 & 0.2 & 0.1 & 0 & 0.7 \end{bmatrix}$$

The experience of observing just one random instance of transmission behaviour for each object in a population with this behaviour can be represented as another matrix, which can be thought of as a random distillation of the transmission matrix in (11):

(12)
$$\begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

(12) represents just five observations: an observation of object 1 signalled by signal 5, an observation of object 2 also triggering signal 5, an observation of object 3 signalled by signal 4, and so on. In the case of object 5, there was no 'choice', as the original probability of signal 1 being transmitted was 1.0, a certainty, In fact, this is the most probable set of observations retrievable from the given transmission matrix. One of the least likely, but still possible, sets of observations from the given transmission behaviour would be:

(13)
$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Such an 'observation-of-transmission' matrix can be generated from an original transmission matrix by making a single weighted random choice of an item in each row. The weighted random procedure makes the probability of a signal being 'observed' for a given object equal to the probability of that signal being transmitted. So the learner's likely experience reflects the statistical tendencies in the behaviour of the population. Where several probabilities in a row are the same, the corresponding signals have an equal chance of being observed in the learner's critical experience of the input adult behaviour. An exactly similar procedure yields an 'observation-of-reception' matrix from a matrix describing the population's mean reception behaviour.

Distilled matrices such as (12) and (13) are taken to represent learners' critical samplings of adult communicative behaviour in the social arena. The learners then have to do something with these samplings to determine their own future behaviour. There are various conceivable simple strategies that might be employed, and I consider three here, which I call the Imitator, Calculator and Saussurean strategies. These names for hypothesized learning strategies are intended to be as apt as possible, within the twin limitations of practical computational simulation and our extensive ignorance about the relevant internal and external constraints on learning signalling systems. I do not suggest that the workings of any organism are likely to be as simple as the definitions given to these strategies, but I hope that they resemble possible real mechanisms closely enough to provide some insight into the ways such mechanisms compete in the evolutionary arena. The definitions of these strategies can be summed up as follows:

(14) Strategy	Sampling of input from adult population		Learner's acquired behaviour
Imitator	Transmission	→	Transmission'
	Reception	→	Reception'
Calculator	Transmission	→	Reception'
	Reception	→	Transmission'
Saussurean	Transmission	→	Transmission' ↓ Reception'

The Imitator strategy is simply to take the samplings of both the transmission and reception behaviour and to make them one's own transmission and reception behaviour, respectively. So the single particular signal that an Imitator observes transmitted for object 1 in a weighted random sampling of the population's transmission behaviour becomes his own constant signal for this object. Likewise with reception; if he observes signal 1 interpreted as, say, object 3, in his random sampling of the group's reception practices, he will in future always 'interpret' signal 1 as object 3. What goes on in the Imitator can be informally summarized (for transmission) as 'Ah, I see someone making signal X in response to object Y, so henceforth I'll make signal X in response to object Y, too'.

The Calculator strategy is to base one's own reception behaviour on one's sampling of the population's transmission behaviour, and vice versa. The way the Calculator derives a transmission or reception matrix from a sampling of reception or transmission behaviour is illustrated below.

$$(13)\quad \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 0.5 & 0 & 0 & 0 & 0.5 \\ 0 & 1 & 0 & 0 & 0 \\ 0.2 & 0.2 & 0.2 & 0.2 & 0.2 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0.2 & 0.2 & 0.2 & 0.2 & 0.2 \\ 0.2 & 0.2 & 0.2 & 0.2 & 0.2 \end{bmatrix}$$

The Calculator's acquired reception behaviour is an optimal response to the observed sampling of transmission behaviour. The derived matrix results from rotating the input matrix (so that rows become columns and columns become rows) and normalizing the resulting rows so that the numbers in them sum to 1, as required of all the matrices we are dealing with. In the derived matrix above, this normalization affects rows 1, 3, 6, and 7. In the case of row (i.e. signal) 1 of the derived matrix, the normalization results from an ambiguity in the transmission sampling, where signal 1 is transmitted for both object 1 and object 5. With a transmission ambiguity affecting two objects (assumed to trigger signalling behaviour with equal frequency), the mean success rate over both objects is 0.5 however one distributes the reception probabilities for that signal between them; it was arbitrarily decided to equalize such reception probabilities. In the case of signals (i.e. rows in the acquired reception matrix) 3, 6, and 7, these signals are actually not transmitted in response to any object in the input transmission sampling; they are 'spare' signals. So whatever reception response one chooses to these signals is in these cases also equally (in)effective. This is possible as there are more potential signals (here seven) than objects (here five), and so with absolute (i.e. probability 1.0) object–signal correspondences, there are always some unused signals. For each ambiguous use of a signal, i.e. use of the same signal for several objects, the number of other unused signals increases.

The Calculator's acquired transmission behaviour is derived in a similar way from a random sampling of the population's reception behaviour, via rotation and possible normalization; his transmission is an optimal response to his sampling of the group's reception.

Neither the Imitator nor the Calculator make any attempt to coordinate *their own* transmission and reception behaviour. The Imitator merely imitates the transmission and reception behaviour of the population as observed in his random sampling; the transmission and reception behaviour of the whole population is not necessarily coordinated, and a random sampling of it is even less guaranteed to be coordinated. But the more coordinated the population's transmission and reception behaviour are, the more successful the Imitator can expect to be, as he will be imitating successful behaviour. The Imitator's strategy is essentially parasitic on developed coordination of transmission and reception behaviour.

The Calculator coordinates his transmission with the population's average reception, and vice versa, in the sense that his behaviour is an optimal response to the population's average behaviour, as seen imperfectly in his sampling. The rotation procedure ensures that the '1's in his sampling of transmission give rise to '1's (or fractions of 1, where there is ambiguity) in just the right entries in his own reception matrix to guarantee him maximum interpretive potential in relation to the transmission sampling. And similarly with his own transmission matrix, designed as an optimal response to the sampling of the population's average reception behaviour. But if the overall population's transmission and reception behaviour are not well coordinated, the Calculator's strategy will not give him individual transmission and reception behaviour that are well coordinated. What goes on in the Calculator can be informally summarized (for reception) as 'Ah, I see someone giving signal X in response to object Y, so I will henceforth interpret signal X as drawing my attention to object Y'; and for transmission, what goes on would be 'Ah, I see someone responding to signal X by attending to object Y, so henceforth I will use signal X to draw attention to object Y'. But in typical cases, this informal story is complicated by ambiguities and synonymies, so that the Calculator has to divide his responses or make arbitrary choices. Either way, communication short of perfection ensues, although the Calculator comes up with the best possible solution available to him.

A plausible real instantiation of the Calculator strategy can be visualized as follows. Imagine a creature capable of learning how to manipulate other creatures, using intentionally produced gestures, and also learning how to respond profitably to actions of theirs. The Calculator treats the problem of how to interact with others without recourse to any notion that he himself may be like them. He learns to interact with them, both transmitting and receiving, as if they were aliens, or of another species. It is as if a person learnt by experience that a certain growl by a lion meant that the lion was about to attack him, and that giving a certain shout tended to drive the lion away. Or, to give an example involving mutually cooperative, rather than adversarial, behaviour, it is as if a person learnt that a certain movement by a cow meant that the cow was ready and willing to be milked, and also learnt that giving a particular gesture of one's own (say holding out a handful of grass) tended to bring the cow near. We all have experience of manipulating and interpreting nature in small ways like this, for example when fishing, hunting, or playing with kittens, or feeding ducks in a park. The Calculator is just an adept learner of such ways of manipulating and interpreting other creatures, in this case creatures of his own kind, who happen to use the same inventory of signals that he himself finds useful.

The Calculator strategy models a capacity to acquire successful goal-oriented behaviour, and its form is not specially tied to symbolic behaviour. It is essentially the ability to acquire optimal interpretive responses to the actions of others, and to learn how to act deliberately on others to best advantage. The Calculator strategy is one strategy, albeit a very simple one, embodying the idea that learning a signalling system is not a special kind of learning, but rather the application of a general intelligent learning ability. In this way, the Calculator strategy contrasts with the Saussurean strategy (below), which models learning in a way consistent with the internalization of a specially linguistic (or more generally, semiotic) construct, the bidirectional Sign. The Calculator strategy would serve well in interactions with creatures whose behaviour is constant and predictable across all their kind. But, as we shall see from simulations, the Calculator strategy is disastrous when deployed in relation to creatures who also learn, that is adapt their behaviour to the behaviour of others.

The Saussurean bases his behaviour on less input data than the Imitator or the Calculator, only taking a sampling of the population's average transmission (and not its reception) behaviour. The Saussurean resembles the Imitator, as far as his acquisition of transmission behaviour is concerned, but he 'takes the law into his own hands' and derives his reception behaviour from his own acquired transmission behaviour such that they are optimally coordinated. Thus, while he can do nothing directly to coordinate the reception and transmission behaviour of other individuals, or even to coordinate his own transmission/reception behaviour with the reception/transmission of others, he 'sets a good example' and makes himself into a microcosm of coordinated transmission and reception behaviour. He makes sure that, whatever his success at receiving the signals of others, he can at least talk to and understand *himself*, as successfully as his transmission behaviour allows. (Incidentally, if language is to be an instrument for facilitating private thought, as well as for communicating publicly with others, it could be necessary that individuals are able to 'communicate' successfully with themselves.)

Informally, what the Saussurean says to himself is 'Ah, I see someone giving signal X for object Y, so X must be the *sign* for Y; I'll accordingly build myself a two-way mapping between X and Y, for use both in transmission and in reception'. The Saussurean's resultant behaviour ends up consistent with an internalized set of signs, each representing a two-way connection between an object and a signal. An incidental consequence of this strategy is that it may produce homonymy, but never synonymy. As learning is based on a transmission sample, with one observation per object, there is never any case of several signals for the same object (synonymy). But there may be cases of homonymy, that is several different objects related to the same signal. As a matter of fact, this parallels the situation in human language, where true synonymy between words is extremely rare, and homonymy quite common.

A large number of simulations were carried out, comparing the relative success of the three learning strategies. The basic assumption was made that mastery of a public communication system increases an individual's reproductive potential. Starting from a variety of initial population behaviours, simulated populations went through successive cycles of (1) birth, (2) acquisition of the basis of signalling behaviour according to the inherited learning strategies just described, (3) selection as parents of the next generation by a random choice weighted according to individuals' communicative/interpretive potential in relation with the population as a whole, and finally (4) death. To tease the reader's curiosity no further, the Saussureans turned out to be clear winners in almost all circumstances; and perhaps surprisingly, the Imitators were markedly more successful than the Calculators, who were the least successful type of all. Further details are given below.

The simulations carried out, and their results, are summarized in table 1. For each position in the table, 20 simulations were carried out; the numbers give the simulations 'won' by each type of learner, within 100 generations. So for instance, all of the 20 3-way simulations (i.e. where all three strategies competed) starting from the 'Random' condition, were won by the Saussurean strategy. Numbers in parentheses give simulations inconclusive after 100 generations.

Table 1.

	Random	Emergent	Perfect
3-way, I v C v S:[a]	20S	17S, 1I (2)	14S, 3I (3)
2-way, I v S:	18S (2)	18S (2)	8S, 6I (6)
2-way, C v S:	20S	20S	20S
2-way, I v C:	17I (3)	20I	20I

[a] I = Imitator, C = Calculator, S = Saussurean

Three starting situations were used, 'Random', 'Emergent', and 'Perfect'. In the Random condition, the initial population had equiprobable responses to all objects and all signals, i.e. no apparent signalling system at all. So, as there were 5 objects involved, the overall population's CP was $^1/_5 = 0.2$. In the Emergent condition, the initial population's average behaviour reflected an emerging connection between some objects and some signals, but by no means perfect communication. In fact, for this condition, the initial population's overall CP was 0.452. In the Perfect condition, a single unambiguous set of object–signal correspondences was reflected absolutely in every individual's transmission and reception behaviour. This population communicated perfectly, with an overall CP of 1.0.

For each starting situation, 20 3-way simulations comparing Imitators, Calculators, and Saussureans were carried out, along with 20 of each of the possible 2-way simulations, Imitator versus Calculator, Imitator versus Saussurean, and Calculator versus Saussurean. The population size was held constant at 30 throughout the simulations. The 3-way simulations started with 10 learners of each kind, and the 2-way simulations started with 15 of each kind of learner involved. These were competitive simulations, carried out to see which strategy turned out to win over its rivals.

In order to force differences between learning strategies to appear within a relatively short time, the relation between communicative/interpretive potential and reproductive potential was deliberately set quite high. In these simulations, an individual's reproductive potential was set proportional to the fifth power of his communicative/interpretive potential. This may well be an exaggeration of the actual selection pressure exerted by expertise in communication, but the technique allows one to get indicative results in a reasonable time. This technique is analogous to that of a physical scientist turning the force of gravity simulated in a centrifuge up to several G. With a weaker relationship between reproductive potential and communicative/interpretive potential, there would have been a larger number of inconclusive simulations and the normal random genetic drift would also have accounted for more results. Nevertheless, the same significant advantage of Saussureans over other types of learners would presumably have been discoverable, but by means of a larger number of longer simulations.

1-way, non-competitive simulations were also carried out, in order to study, for the different strategies, the rate at which uniform populations converged on coordinated signalling behaviour and at what level of communicative competence.

5. The trouble with Calculators

Why do Calculators fare so badly, when their learning strategy is calculated to yield optimal transmission and reception responses to samplings of the reception and transmission behaviour of the whole population? The answer lies largely in the degree of indeterminacy involved in deriving the Calculator's behaviour from that of the model population behaviour. The derivation of learner behaviour from population behaviour involves a double indeterminacy for an Imitator, a different kind of double indeterminacy for a Saussurean, and a quadruple indeterminacy for a Calculator. I explain this below.

All strategies, Imitator, Calculator, and Saussurean, suffer from the indeterminacy involved in sampling. If the whole population's behaviour is not completely uniform, then separate samplings may yield different inputs for learning. For instance, if some individuals give signal 1 for object 1 and others give signal 7 for this object, some transmission samplings will show a 1–1 object signal mapping and others will show a 1–7 mapping, so all learners in a heterogeneous population tend to suffer from the atypicality of their samplings. One single description of a whole population's behaviour gives rise to many different possible samplings.

(16)

$$\text{Population behaviour} \rightarrow \begin{cases} \text{sampling 1} \\ \text{sampling 2} \\ \vdots \\ \text{sampling } n \end{cases}$$

This indeterminacy affects transmission and reception behaviour equally. The samplings used in simulations only took one 'look' at each row of the population's mean transmission and reception matrices, and so, with quite heterogeneous behaviour at large in the population, the effect of the sampling indeterminacy can be quite marked. But it is a general fact about sampling, in whatever domain, that it involves indeterminacy. Indeterminacy can be reduced by more extensive or more systematic sampling methods, but it can never be eliminated.

For the Imitator, this sampling indeterminacy affects the acquisition of both his transmission and his reception behaviour, but there are no further indeterminacies in the process of acquiring behaviour patterns of his own. However nothing in the Imitator strategy relates to the coordination of transmission and reception behaviour.

The Saussurean resembles the Imitator as far as his transmission behaviour is concerned, but acquires his reception behaviour by a process which replaces the sampling indeterminacy with a different kind of indeterminacy. The Saussurean's reception behaviour is an optimal response to his own acquired transmission behaviour. For some, indeed most, transmission samplings, there can be several equally optimal responses to them. For instance, where the same signal (say signal 7) is observed transmitted twice, once each for two different objects (say object 1 and object 2), an optimal response will interpret this signal as either object 1 invariably or object 2 invariably, or sometimes object 1 and sometimes object 2; these different responses are all equal best. There are many different optimal responses to any one particular sampling.

(17)

$$\text{Sampling} \rightarrow \begin{cases} \text{optimal response 1} \\ \text{optimal response 2} \\ \vdots \\ \text{optimal response } n \end{cases}$$

This is equally true of both transmission and reception behaviour, but the Saussurean only takes this step in relation to one side (the reception side) of his behaviour. Thus the Saussurean's acquisition of reception behaviour involves him in a different indeterminacy from the sampling indeterminacy, but insofar as this arises from an attempt to optimize his reception in relation to his own transmission, an element of coordination between transmission and reception is introduced into any population to which he belongs.

The Calculator suffers from the sampling indeterminacy, for both transmission and reception, just as does the Imitator, but also gets a double dose of the optimal response indeterminacy, as he uses the optimal response tactic to derive both his transmission and his reception behaviour from samplings of the population's behaviour. And, although his acquired behaviours are optimal responses to the population's, there is no element of coordination internal to the Calculator himself; he is not constrained to be able to communicate well with himself.

The comparison between strategies can helpfully be summarized somewhat formally. Say the number of possible samplings of transmission behaviour from a heterogeneous population is t, and the number of possible reception samplings r. Say also that the number of possible optimal responses to a sampling is given by the function Opt. Then the numbers of possible different individual behaviour patterns that can be acquired from such a population via the different strategies are as follows:

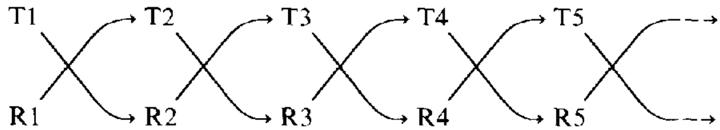
(18) Imitator:	t	x	r	
Calculator:	$Opt(r)$		x	$Opt(t)$
Saussurean:	t	x	$Opt(1)$	

This shows that in terms of computational complexity, the Imitator's strategy is the simplest of the three, and the Calculator's strategy is the most complex, with the Saussurean's strategy intermediate in complexity between the other two.

Clearly, the number of different behaviour patterns accessible via the Calculator strategy is greater than those accessible via the other two strategies. Without considering the selective advantages of better communication, this means that, in a population of all Calculators, where a group signalling system is passed from one generation to the next via the Calculator strategy, the behaviours of successive generations will tend to become increasingly diverse, up to some limit. In a population of Saussureans, with no effect of selective advantage, there would also be some diversification, but less markedly. In a mixed population of Calculators and Saussureans, presumably the Saussureans in successive generations would tend to reflect core group behaviour, whereas the range of Calculator behaviours would tend to be greater. This would tend to give individual Saussureans greater communicative potential, in relation to the rest of the population, than individual Calculators.

The high degree of indeterminacy involved in learning via the Calculator strategy can be seen in an illustrative computation. An initial pair of transmission and reception matrices for a hypothetical Calculator was set up, reflecting a perfect 1–1 correspondence between 5 objects and 5 signals (with 2 signals 'spare', i.e. having zeroes in their columns in the transmission matrix, and showing equiprobable responses (0.2) in all entries of their rows in the reception matrix). From each of these matrices a corresponding new reception and transmission matrix, respectively, was derived, via the sampling and optimal response procedures described. And these 2 matrices were in turn used to derive a 3rd pair of transmission and reception matrices, and so on and so on. This can be taken to represent a continuing population consisting of a single individual Calculator per generation, with each generation acquiring its transmission and reception behaviour via the Calculator strategy by sampling the reception and transmission behaviour of the previous generation and responding optimally to it, as shown in (19).

(19) A one-Calculator population through time



In this experiment, the original situation of perfect communication was immediately lost and never regained. After the initial T1 and R1, subsequent T*n* and R*n* resembled a random walk around the space of possible T and R matrices. In 14 iterations of the process, no matrix ever recurred. By contrast, when corresponding experiments were done with learning mediated by the Imitator and Saussurean strategies, the original situation of perfect communication was preserved intact at each stage. In these cases, the corresponding diagrams look as follows:

(20) A one-Imitator population through time

T1 → T2 → T3 → T4 → T5 - - →
R1 → R2 → R3 → R4 → R5 - - →

(21) A one-Saussurean population through time

$$\begin{array}{cccccccccc} T1 & \rightarrow & T2 & \rightarrow & T3 & \rightarrow & T4 & \rightarrow & T5 & \dashrightarrow \\ & & \downarrow & & \downarrow & & \downarrow & & \downarrow & \\ R1 & & R2 & & R3 & & R4 & & R5 & \dashrightarrow \end{array}$$

A clear demonstration of the ineffectiveness of the Calculator strategy was obtained by a series of simulations starting with a homogeneous population of 30 Calculators with perfectly coordinated transmission and reception matrices. All individuals in this initial population communicated perfectly with all others, about five objects, using seven possible signals; thus the CP of this initial population was 1.0. The population bred at random, with probability of being a parent heavily weighted in favour of those who communicated best with their fellows. (At first, of course, there was no such difference between individuals, as they all communicated perfectly.) New individuals in each generation acquired their communicative behaviour, via the Calculator strategy, from the average of the behaviour of the previous generation. The overall CP for the population was measured at each generation. The rapid decline to a level of mere chance success is shown in figure 1.

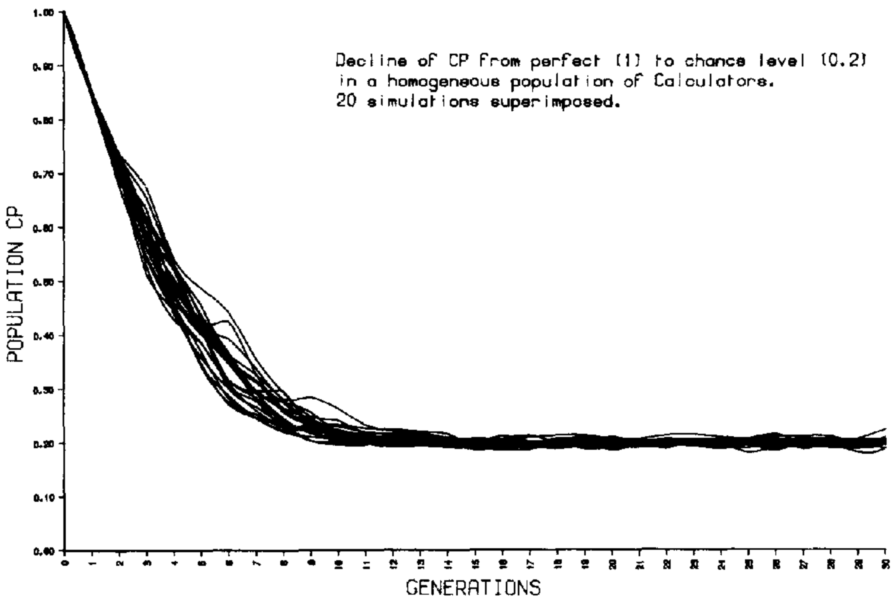


Fig. 1.

Figure 1 can be said to show graphically that Calculators are, colloquially speaking, too calculating for their own good. The situation is roughly analogous to one in which a well-intentioned helper at a rifle range moves the target in a calculated way to compensate for a marksman's misses, while the marksman himself also adjusts his own aim, assuming that the target will remain where it last was. Too many people calculating, like too many cooks, leads to chaos. Someone needs to just stand still, and let others calculate their way to him.

Figure 1 can be contrasted with figures 2 and 3, which give the results of similar simulations carried out with populations of Imitators and Saussureans. These simulations started with homogeneous populations of 30 Imitators (or Saussureans) with random transmission and reception matrices, assuming five objects and seven possible signals. That is, in these initial populations, no individual showed any bias towards any particular signal for any object, or any particular object as the interpretation of any signal. Thus, as each individual communicated with only a chance level of success with all others, the CP of these initial populations was 0.2. As before, the populations bred randomly, with the privilege of parenthood tending heavily to be allocated to the better communicators. New individuals in each generation acquired their communicative behaviour, via the Imitator or Saussurean strategy, from the average of the behaviour of the previous generation.

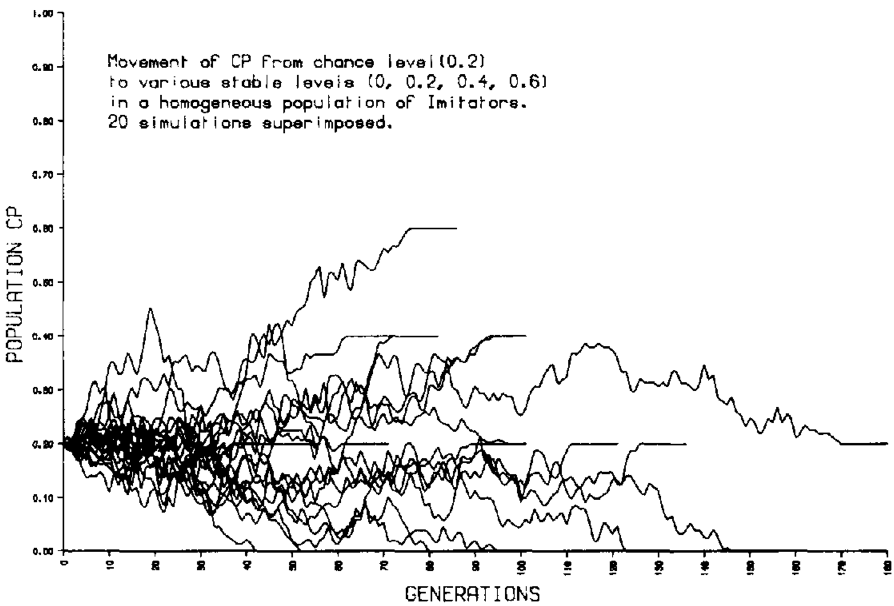


Fig. 2.

For the Imitator strategy, the population's overall CP evolved in 20 trials to the following levels:

(22)	0 (i.e. consistent failure of communication)	in 6 cases
	0.2 (success on 1 object, the chance level of performance)	in 9 cases
	0.4 (success in relation to 2 objects)	in 4 cases
	0.6 (success on 3 objects)	in 1 case

Hence, whereas a population of Calculators, starting from any level of communicative success, even from perfect communication, will almost inevitably decline quickly to mere random performance, some populations of Imitators, starting from the random level, can happen by chance to evolve to more successful levels, although in the majority of cases they end up at or even below the random level. Naturally, if a population of Imitators is 'given a better chance' and starts off at a better than chance level of communicative success, it will tend to evolve to higher levels than otherwise. And if it is started at the level of perfect communication (CP = 1), a population of Imatators will in all cases actually preserve this perfect communicative situation through successive generations; this follows from the nature of the Imitator strategy for the acquisition of communicative behaviour.

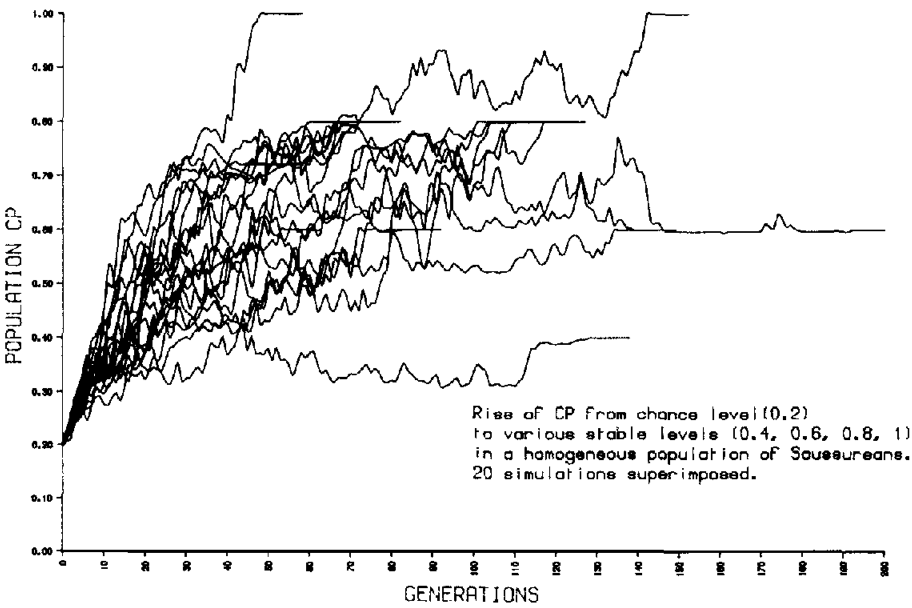


Fig. 3.

For the Saussurean strategy, the population's overall CP evolved in 20 trials to the following levels:

(23) 0.4 (success in relation to 2 objects)	in 1 case
0.6 (success on 3 objects)	in 6 cases
0.8 (success on 4 objects)	in 11 cases
1.0 (perfect communication)	in 2 cases

Clearly, the Saussurean strategy is more successful than either the Imitator or the Calculator strategy. As with the Imitator strategy, if a population of Saussureans starts off at a better than chance level of communicative success, it will tend to evolve to higher levels. And if started at the level of perfect communication (CP = 1), a population of Saussureans will always, like a population of Imitators, preserve this perfect communicative situation through successive generations.

Obviously, all the computations described above are extremely idealized, and the three strategies compared almost certainly don't exist in these stark forms in nature. But I do not believe that I have loaded the dice by idealizing any of these strategies in such a way as to render it less (or more) successful. It appears that a genetically transmitted acquisition strategy for simple communication systems that incorporates the essence of the bidirectional Saussurean sign has very clear advantages over other conceivable strategies. Although much remains to be debated, I claim that mechanisms of the sort described are plausible idealized reconstructions of what took place in the earliest stages of the long evolutionary road to the present species, with its capacity to acquire, and use for communication, the richly structured systems of today's languages.

References

Clutton-Brock, T.H. and P.H. Harvey (eds.), 1978. Readings in sociobiology. San Francisco, CA: Freeman.

Dawkins, R., 1986. The blind watchmaker. Harlow, Essex: Longman.

Hamilton, W.D., 1971. Geometry for the selfish herd. Journal of Theoretical Biology 31. Reprinted in: T.H. Clutton-Brock and P.H. Harvey (eds.), Readings in sociobiology, 142–159. San Francisco, CA: Freeman.

Hurford, J.R., 1987. Language and number. Oxford: Blackwell.

Hurford, J.R., to appear. Pre-puberty, the critical period for language acquisition: An evolutionary explanation.

Kitcher, P., 1985. Vaulting ambition. Cambridge, MA: MIT Press.

KCSG (= King's College sociobiology group, Cambridge) (eds.), 1982. Current problems in sociobiology. Cambridge: Cambridge University Press.

Maynard Smith, J., 1965. The evolution of alarm calls. American Naturalist 99. Reprinted in: T.H. Clutton-Brock and P.H. Harvey, (eds.), Readings in sociobiology, 183–188. San Francisco, CA: Freeman.

Maynard Smith, J., 1976 Evolution and the theory of games.. American Scientist 64(1): 41–45. Reprinted in: T.H. Clutton-Brock and P.H. Harvey (eds.), Readings in sociobiology, 258–270. San Francisco, CA: Freeman.

Orians, G.H., 1969. On the evolution of mating systems in birds and mammals. American Naturalist 103. Reprinted in: T.H. Clutton-Brock and P.H. Harvey (eds.), Readings in sociobiology, 115–132. San Francisco, CA: Freeman.

Reynolds, V., 1980. The biology of human action. San Francisco, CA: Freeman.

Rose, S., R.C. Lewontin and L.J. Kamin, 1984. Not in our genes: Biology, ideology, and human nature. Harmondsworth, Middlesex: Penguin Books.

Saussure, F. de, 1959. Course in general linguistics (transl. Wade Baskin). New York: McGraw-Hill.

Sperber, D. and D. Wilson, 1986. Relevance. Oxford: Blackwell.

Trivers, R.L., 1971. The evolution of reciprocal altruism. Quarterly Review of Biology 46, 35–57.