

## How language works

*D. Crystal*

December 14-16, 2008

Crystal is an untiring popularizer who has produced a steady stream of books and even TV-documentaries addressing the educated public on the topic of language, in particular its English manifestation, the mother tongue of the majority of his reader. The present book is a systematic compilation of what constitute linguistics, a field that does not cease to fascinate the author, ranging widely as it does touching on physics (the formation of sounds), physiology, neurology, history, language structure and language diversity, education, and perhaps most intriguingly on philosophy and the existential question of what it means to be human, in which language maybe the single-most important defining feature. Of course the ultimate questions on language are of course those that fuel our fascination and gives it a direction and a content, but which cannot be properly addressed and formulated, unless there is a thorough empirical grounding.

Language is a biologically induced phenomenon (and as the evolutionist Maynard-Smith claims, one of three great innovations of evolution alongside with multi-cellular organisms and sexual reproduction) but once launched transcending its fortuitous origins. As a biological phenomena it is about speech, writing being a later cultural embellishment<sup>1</sup>. Speech is about forming sounds rapidly and understanding them in real time, a process of remarkable neurological agility, in many ways similar to muscular control, and indeed as to the forming of speech, as opposed to interpreting it, it is all about muscular activity controlling the movements of tongue and lips and the flow of air. The human anatomy is very well adapted to the generation of vocal sounds, this being one of the main anatomical differences between us and our closest cousins, the primates<sup>2</sup>. Incidentally one of the disadvantages of such an ability is the possibility to choke on food, the wind tracts having to perform dual roles, a calamity that does not befall animals (and very young infants). Depending on the position of the tongue and the flow of air, a great variety of various sounds can be produced voluntarily, only a fraction of which is used in anyone language<sup>3</sup>. Such units of vocal sound are referred to a s phonemes, and they can be classified by a few parameters, pertaining to position of lips and tongues and the art of the flow and such things, making up in fact a continuous configuration space (varying from individual to individual, and strictly speaking not being individually invariant either but changing ever so slowly with time) which we more or less arbitrarily have divided up discretely. The

---

<sup>1</sup> As Dawkins, the evolutionary prophet par excellence, no doubt would characterize as the launching pad for a new kind of evolution in the sense of being non-biological, 'memes' instead of genes

<sup>2</sup> It is an unresolved issue to what extent the extinct Neanderthal man had a vocal ability, it is assumed that he had at least a rudimentary one, being able to produce some notable fraction of the possibilities available to modern man

<sup>3</sup> The babbling child is capable of reproducing them all, the acquisition of a mother tongue meaning a particular restriction of them as it linguistically matures

number of phonemes employed in a language is often significantly more than the number of letters used to assign them, yet not larger than making the task of devising a comprehensive phonetic alphabet very feasible<sup>4</sup>. The result is of course a confusion in spelling, supposedly to be particularly egregious in English<sup>5</sup>. Now phonemes are sounded differently depending on where in a word they appear, being both affected by what has come before and what is anticipated to follow; thus the notion of a perfectly spelled language is something of a myth.

Phonemes combine to form syllables, sometimes only consisting of one phoneme, and syllables form morphemes, the smallest unit that makes any sense (if only grammatically). Words themselves are formed out of morphemes<sup>6</sup>. Words are the basic central units in a language, and just as phonemes are given by nature (i.e. anatomically restricted), words themselves form a fixed (but rapidly evolving) collection determined by convention, not to be tampered with by the user<sup>7</sup>. Now words can be combined into sentences, building up to more or less extensive texts, and here the combinations are virtually endless, limited only by the imagination of the speaker, as well as certain syntactic rules (mostly unconscious ones). While the forming of new words is usually looked down upon as a sign of ignorance or cheating (or both), the failure to produce new texts is looked upon as close to criminal and referred to as plagiarism.

The above is a thumbnail sketch of the formal structure of language. As such it is in many ways misleading, even if instructive and inspirational, being an artifact of the written encoding of language. The notion of parsing speech into words (normally delineated by spaces) is not clear. Compound words, such as in German, are considered to be single words by virtue of being written without spatial interruption and being referred to as entities. But the user of German has a great liberty in forming new compounds, and such are not normally dismissed as neologisms, akin as they are to the forming of sentences. Also, as Collingwood remarks, idiomatic phrases really should be considered as words themselves, and are in fact part of the basic vocabulary a user steadily accumulates<sup>8</sup>. The formal structure of language has inspired the creation of artificial languages, those involve alphabets, and a subset of combinations of those, referred to as the vocabulary (which by the process of 'macros' can be indefinitely extended), elements of which will form sentences according to strict syntactic rules. An artificial language could on one hand be studied

---

<sup>4</sup> Which has in fact been done repeatedly. A simplified version of it I encountered for the first time when I started to study English at the age of eleven. That was of course logical, yet in retrospect I suspect it could have been an additional burden to many new learners.

<sup>5</sup> The erratic spelling of English is to a large extent a myth, meaning meaning both in general that idiosyncrasy in spelling is much less of a hurdle than what one would naively expect, and that the amount of irregular spelling in English is in fact far less prevalent than one would at first suspect.

<sup>6</sup> Just as a syllable can consist of only one phoneme, a morpheme can consist of only one syllable, and a word of a single morpheme, hence the existence of monophonemic words in many languages such as *I,a* in English and *i,,* in Swedish

<sup>7</sup> neologisms, just as mutations, are frowned upon (Defoe even likened the coining of new words to be akin to the forgery of coins), but without their intermittent occurrence, languages (just as life) would not evolve, nor develop.

<sup>8</sup> there is a thin line between the idiomatic expression and the so called hackneyed phrase

just as a formal entity generating all kinds of intrinsic questions, or be made to refer to beyond itself, such as in computer languages, where a given text - a program, can be made to interact with reality, i.e. some computer hardware and its extensions, and cause all kinds of real life phenomena (such as the extermination of civilization by a nuclear holocaust). However artificial languages are not real human languages, in particular they do not allow themselves to talk about themselves in any but the most frivolous of senses. Those considerations do not belong to linguistic proper, being too metaphysical in nature, but to its philosophical aspects, probably the most interesting aspects of language, and to which we will ultimately return.

In order to understand language you need to study how it is acquired. The basics of language is mastered by the growing child in just a few years. The prerequisite for such a mastery is to understand what the point of language is, just as the first step in trying to decode a message is to understand that it is a message in the first place (also assuming the understanding of the notion of a message, but that understanding might conceivably be concomitant with the previous). Thus an extremely important stage in the linguistic development of a child is to get attuned to the notion of communication, in particular as conveyed by language. This is done during the first year when it learns to pay attention to voices and that those can be significant. It is notable that we can all distinguish human vocal intercourse as such, even if we do not understand the language employed at all. Speech is often compared to music, sometimes the two are literally conflated as in singing, but there is a fundamental difference. Music is based on repetition of themes and almost symmetries, the sounds by themselves do not make up meaningful entities, their effects are made through contrasts. Speech on the other hand calls attention to its meaningful entities out of which a message is built, intrinsic audial features are generally unnoticed, except possibly to some limited extent in poetry, where the sound is but a complement to the sense. It is true though that a language we do not understand, or so poorly that we are unable to construct the meaning when we hear it, we can judge externally by aesthetic criteria, something which is impossible for a language with which we are genuinely intimate<sup>9</sup>. We do not learn to appreciate music by splitting it up in basic sounds and learning how to combine them, we learn to savour it by taking it in as a whole. Admittedly something similar may be the case with the learning of language, to which we will return. Now spoken language carry with it many additional features wholly absent in writing. The tone of voice can convey important messages, universal in their appeal, even comprehensible to fellow mammals<sup>10</sup>. This is why written communication, even by the best of writers, may result in grave misunderstandings (especially the simplified e-mail message), misunderstandings that would never arisen in normal face to face contact, and often can only be resolved (if at all) through such direct procedures. In particular irony is very liable to get lost in the

---

<sup>9</sup> Is your own mother tongue beautiful as a sequence of sounds? We cannot tell, the meaning of any discourse intrudes itself and makes such an assessment impossible. Possibly one way of getting around it would be to play it backwards.

<sup>10</sup> It are those extra features of spoken language which makes it possible to communicate with pets. They may not understand the words, in fact probably having no notion of words, being alingual by nature, but they may nevertheless sense the intention behind an utterance, whether it is one of affection or anger, and react accordingly.

written word, to nevertheless sense it in a written text inspires in the reader a conviction that there is a special bond between him and the writer, excluding the rest. This is why students of Plato often exalt him, and which can to some extent explain the tone deafness of a Popper (but this is an illegitimate digression.)

After the child has been socially invited into the sharing of language, it acquires it with an amazing speed. The ease with which children learn their mother tongue, and the near universality of the success, makes it very reasonable to assume that the human child is prewired for the task, just as he or she are prewired to learn to walk. In short children learn to talk by instinct. This is extremely interesting as to its implications to formal education. Could it be that most things we learn, we learn either by compelling instinct, or through necessity, and that the formal schooling that is more and more available to the populace, contributes but a small fraction of the expertise of the ordinary individual, and that it is only really beneficial to a small section of the population, who, for some reason or another, have learned to become predisposed to its fruits? This is of course another digression to be abandoned, and let us return to the main theme. What is interesting about the learning process of a child is to note the errors. The errors can be intriguing windows into the deeper structure of language, but the problem is that the errors are very few and give but brief glimpses of what lies below. It is true there is a growing complications of the child's utterances, as one would expect from mere armchair speculations. The child at first only utters single words, not because of the simplicity of its thoughts (and the relation between thought and verbal expressions is from a linguistic point of view metaphysical and the discussion will hence be postponed) but because of its limited verbal command. That single word, say 'Daddy' will have to convey a variety of meanings, distinguished not only by tones of voice and gestures, but the given context. As the child grows as to verbal sophistication, it will combine words, first one and one to two, but gradually making clusters, and super-clusters, grammatically referred to as clauses. Yet in any communication, the context, whether verbal or not, cannot be ignored, no matter the level of verbal sophistication. This is something which is generally not appreciated, because the context is usually so basic that it becomes invisible.

The child thinks and it needs to communicate its thoughts. The desire to communicate is frustrating and presents endless problems to be solved. And as Popper indicates, what has relevance to the solution of a problem that stymies us, will invariably grab our attention and be attended to and assimilated. That the child acquires a vocabulary so quickly and retains it is remarkable but not really miraculous. Words are often related to each other, and the vocabulary of an individual, is not a long random list to which acquisitions are tagged onto its end, but it is a structured one, in which words are linked, and new words are placed into a context<sup>11</sup> of linked ones. What maybe noteworthy though is that it seems much easier for a child to learn thousands of spoken words than to learn visual characters for those. The former being in the nature of a natural neurological process, the latter

---

<sup>11</sup> Crystal makes (a probably useful) distinction between the linking of related words, such as big-small, which is intra-subjective, and the purely subjective associations between words, which make up the personal history of an individual, and which psycho-analysts of the past, saw as a possible gateway into the mind of a patient.

apparently an awkward culturally imposed one<sup>12</sup>. What is truly miraculous though is the acquisition of syntax, there are so many ways of constructing ungrammatical sentences (including those that adhere to all explicit grammatical rules, but still do not 'sound right') but few of them are ever stumbled across. As Chomsky has noted, there is simply too little input for the child to empirically base its expertise, word patterns simply have to be innate. Yet to anchor those speculations into neurological terms, or to articulate them through a concrete testable theory, is quite another thing. Suggestive as those speculations may be, and I certainly feel that they have the ring of truth, it seems very hard to formulate them into explicit generative rules. According to Chomsky there is a universal grammar, out of which all the grammars of all human languages are just particular manifestations given a set of parameters (initial conditions). But how should one think of such a grammar? Maybe one would need a very wide as well as penetrating acquaintance with a few hundred widely scattered languages, which is beyond the capability of a single individual, in order to get a clue?

How is language really acquired? By instruction or osmosis? There is a fundamental difference between how a first language is made possession of and how a second is acquired. The first is a momentous experience and serves as a template to the second, in the process contorting it<sup>13</sup>. Many features, such as a highly inflected language, that present a formidable obstacle to the instructed learner, is taken in stride by the instinctive one. The reason being that the uninstructed learning of language does not involve explicit categories. Words are not learned in isolations, nor grammatical features. Words always figure in sentences, which are laden with meaning. The uninstructed language learner probably does not realize that there are things like words, just as we are not aware of how we place our tongue and lips when we produce sounds, unless we are explicitly instructed<sup>14</sup>. Thus a word and all its inflections are derived from contexts, top-down rather than bottom-up. It would be naive to conclude from this that there is a natural method of learning foreign

---

<sup>12</sup> The spelling of words, presupposing an alphabetic, or at least syllabic script, is just a decoding of information already acquired; but the association of words to ideograms is a new process. The Chinese and Japanese student spend many years just in mastering a few thousand basic characters, which is a prerequisite for basic literacy. One may argue that when once mastered, this enables the student to more quickly peruse a text, than what a reader of a script is able to do. One should keep in mind though that the accomplished reader of a script, does not painstakingly decode the meaning of a word, common words at least are instantly recognized by their appearance, not unlike that of familiar characters, others are expected and confirmed by checking a few salient features. This is why being a poor speller usually does not impede reading. Why bother about 'occassion' versus 'occasion' (or even 'occacion') both look essentially the same? On the other hand we seldom get confused between 'insect' and 'incest' although they only differ by a seemingly innocuous transposition of two letters. One can for a lifetime of reading be unaware of that extra 'u' in 'continuous' and consistently spell it as 'continous' even if one encounters it hundreds of times each day, because both spellings would pass the usual tests we employ when we expect the word to occur.

<sup>13</sup> This discussion ignores how two languages can be imposed simultaneously on the growing child, one would expect that one will take preference, but the preferred language can differ depending on context

<sup>14</sup> and then this does not necessarily help matters, on the contrary, as anyone subjected to logopedic instruction can testify to

languages, and that furthermore this natural method would be a reconstruction of the instinctive one, as it was once the vogue to believe. The educator faces a dilemma, maybe an insurmountable one. By definition instincts are instincts and cannot be replaced by instructions, any attempt at instruction has to employ some explicit stratagems (such as our formal description of language) which goes counter to any instinctive process. Also, it could be the case that once a language is acquired, the process cannot be repeated, innocence has been lost once there is a language in residence, it invariably affects neurological pathways<sup>15</sup> .

Now as a user becomes more adept at using (and abusing) a language, there is a convergence between him and the native expert, as to the acquisitions of new vocabulary<sup>16</sup> or picking up stylistic nuances. It is not unusual for a foreigner to acquire a mastery of the adopted tongue denied the majority of its native speakers, this in particular holds for the educated. The really interesting thing is whether this is a formal mastery, and that the foreign speaker will for ever be barred from a complete mastery, either through the inability to have a complete idiomatic command, or to be liable to occasionally slip into mistakes, a native one would be unable to make, however uneducated. When it comes to accent, this seems to be the case, it is sometimes claimed that one may never be able to learn to speak a foreign language without an accent, unless one learns it before puberty. For the sophisticated speaker, this particular issue is of minor concern, he or she would be more concerned with the ability of expression at a literary level. Few writers make a career of writing in a foreign tongue, conventional wisdom holds that it is simple impossible, yet there are exceptions, Conrad being one of the more illustrious ones, and consequently his accomplishment has been met with some disbelief, and there has been a search for infelicities of his English in his work. On the other hand a truly great writer is supposed to further a language, to renew it, press it beyond what it is expected to perform. A truly great writer has to produce the occasional infelicity otherwise he will have had no influence on the language at all.

The most fascinating aspects of language to the public, is its diversity. We are speaking not only about the diversity of language worldwide, but maybe even more about internal diversity, i.e. dialects of a single one, most commonly of that of the own. Connected to this interest is the idea that dialects are but unsophisticated distortions of the pure language, in particular as it is encoded in written form. Few things rouse the passion of people, or at least some people, as the perceived abuses of language, set against some supposed norms. Crystal has little patience for such people and their concern. For one thing it is hard to make a distinction between languages and dialects, even harder than to make a distinction between species and races. In the latter case one traditionally uses the ability to sire fertile descendants, and similarly the criterion of mutual intelligibility (without former instruction and training). But such criteria face problems with transitivity, making the

---

<sup>15</sup> One may think that the way to overcome this would be to erase the resident mother tongue, how this could be controllably effected without further destruction seems a very difficult matter, and hope that the instincts for language would be unaffected. Then in the end what would be the point?

<sup>16</sup> Most words you know in a foreign language have not been learned consciously by looking them up, but my natural assimilation from the context.

partition into disjoint classes impossible<sup>17</sup>. The traditional quip by the linguist is that a language is a dialect with a navy. Thus we are speaking about conventions, yet hard as it is to make precise boundaries and definitions, seen from a high perspective, the rough classification into different languages is very useful.

Dialects of your own language form a topic of interest, because all speakers have some expertise on the matter and are daily confronted with the phenomenon, thus able to make their own assessments and form their own strong opinions. Language diversity is but dialectical diversity at its extreme. This has spawned the notion of a family tree, that languages are related to each other through the spawning of descendants. This has attracted a lot of forensic philological interest and extinct parent languages have been reconstructed through the fossils they have dispersed through their descendants. The investigation of the history of language was started at the end of the 18th century by the observation of William Jones<sup>18</sup> to the effect that Sanskrit and its modern vernacular versions showed remarkable similarity with western European languages. Similarities which were so extensive and systematic that they simply could not be accidental. Concomitant with the interest in the origins of languages was an interest into racial history. However, race and language have nothing to do with each other. Language as such may be a biological feature, the way it is manifested is a cultural historical accident. It is remarkable indeed, that any child irrespective of racial make-up, can become a native speaker in any language, provided that he or she is exposed to that particular language during their formative period. Indeed the politically correct opinion of our times that we are all humans, that racial diversity is but accidental and has no bearing on our humanity seems to be born out of biology<sup>19</sup>. Thus languages spread like waves do, because just as in the case of waves the medium does not move, only what is manifested in the medium. This should be obvious (say in view of all the Black Americans whose native tongue is English), but remarkably the myth is well entrenched. Thus in particular there is little to support the thesis of a radiating migration of Aryan people from Central Asia. The majority of the ancestors of the Scandinavian population may very well have spoken some Finnish type language and the Germanic tongue may have been an outside influence. This is of course all speculation.

The diversity of tongues world-wide is, pushing the biological metaphor further, a source of wonder and fascination valuable by itself, just as biological diversity. But just as with most of the interesting species, most languages are at risk for extinction. Lan-

---

<sup>17</sup> As to races there is the phenomenon of monodromy, sub-populations of arctic terns, interbreeding in overlaps as we travel east, but once the pole has been circumscribed, interbreeding is no longer possible. Similarly there is a continuum of Germanic dialects across Northern Europe, starting with Dutch imperceptibly flowing into low German then down into the dialects of Austria and Switzerland, locally mutually intelligible, but not as respect to its extremes.

<sup>18</sup> An British Civil Servant stationed in India acting as a judge. Typically the first European visitors to India were fascinated by its culture and often took as a result of this a keen interest in it, in particular to probe into the local languages.

<sup>19</sup> This is to some extent fortuitous, biology has in the past been harnessed to give support to the opinion of racial hierarchies, but apparent initial success notwithstanding, that success was spurious, and on the contrary biology has found nothing to support such claims. The situation would potentially have been far more morally ambiguous had Neanderthal men survived into the modern age.

guages are social phenomena, and once the basis for them are being eroded they become superfluous and eventually obsolete. The great majority of languages are only spoken by a few thousand people, and such numbers do not provide critical mass. Speakers of such languages cannot afford to be monolingual and to survive in a large world they do need to speak at least some dominant lingua franca, in addition to a few more neighboring tongues. Sometimes if many different speakers are thrown together a new common tongue has to be devised. a makeshift construction to serve the immediate needs. Such languages are indeed quite primitive, yet linguists claim that they display some common features which are biological and not historical. Thus in this view so called pidginization is a natural phenomenon appearing spontaneously throughout human history. Pidgins are not fully developed languages, but once they became option to become first languages to hapless children, the latter turn them into real languages. This process is referred to as creolization. This points to a significant complication when it comes to the idea of a language tree. Languages, unlike species, can occur not just by diversification, but by joining unrelated languages (a milder version of this is seen in the propensity of languages to borrow vocabularies from each other, English being a particular eager gobbler of that kind). Thus the simple tree-like structure is misleading. In fact there are many ways of classifying languages, and indeed relatively distant languages may share features with each other that they do not share with closer relatives<sup>20</sup>. The ultimate question of language history is whether there were at some time one single language that has become diversified, and if so when this single language started. More specifically one would like to find dates in the past when the different Germanic or Slavic languages started to diverge. Based on some general evolutionary principles of neutrality, such datings have been achieved for the bifurcations of species, due to molecular drift, it is unlikely that something similar can be achieved for languages. In fact the notion of a universal parent tongue seems naive. The human race did not start with an Adam and Eve speaking a primordial tongue, the roots of language diversity have to be looked for in pre-human history.

Can you have many mother tongues? Or once having acquired a first language can you ever hope to acquire another language in which you can achieve the same mastery as with the first? Does there exist such a thing as true bilinguality? The consensus seems to be that this is truly rare, one language invariably takes precedence over another. Still if the two languages are very close, say one an informal dialect of the other, the distinction may be moot as to which really is the primary language. Other spurious instances of multilingualism are due to the fact that different walks of life require different vocabularies. Those may due to personal life experience belong to different languages, and as a consequence no language seems to take precedence over the other. Some activities, like a certain professional life may by preference be conducted in one language, while other activities would be more comfortable in another, due to the availability of pertinent vocabulary. Yet, when it comes to deeper structures, one language invariably would tend to be dominant. Then there is the question of mastery. Traditionally an educated Scandinavian was expected to be versant in three languages in addition to his Native tongue, namely German, English

---

<sup>20</sup> One example of this is the practice of having the definite article transform itself to an ending of the noun. This is the case of Scandinavian languages, Rumanian and Bulgarian, but not shared by other Germanic, Romance and Slavic respectively.



and French (just as an educated Anglo-Saxon a hundred years ago, could include the former in the latter). Mastery obviously varied between the languages and individuals, and some may have in addition to the requirements added a smattering of two or three other languages. In the past multilinguality also involved dead languages such as Latin, Greek and Hebrew (subsequently revived), and there are stories of linguistic prodigies that by the age of five or so, supposedly mastered five classical tongues. How many languages can a human absorb? There are stories, maybe apocryphal, about some individuals mastering up to fifty different languages. What is meant by mastering? It is said that once you know three or four languages, the addition of new ones is comparatively easy. This may be true, but it still does not address the basic question what is meant by true mastery? To limp along with a tongue is one thing, and here I would suspect that there is no restrictive limit, but what about being able to make jokes, being sarcastic, know when you are being rude to. I would suspect that with each language you truly have a command of, there goes a corresponding personality, and how many different personalities can you have, and still be an integral being? But a truly multilingual individual, who did master say a variety of widely different languages (say Chinese, English, some North American Native tongue etc), would have a unique stereoscopic vision, being able to ascertain what is merely accidental in each particular language manifestation, and what is common to all human linguistic expression. And also to be able to confirm, what is suggested, that all mature languages are equally expressive and syntactically sophisticated. But do such individuals exist? The science of linguistic can at least effect comparisons, but of that fascinating exercise there is preciously little to be gained from the book.

Finally language is ultimately about meaning. Verbal expressions are made with intentions and convey meanings, which are somehow transmitted and intuited. This is the meta-physics of language, beyond the formal description. But is meaning an emergent feature, and the sounds of speech precede the meaning they ultimately are called upon to serve? Or does thought exist beforehand desperately trying to manifest itself ever ready to avail itself of any opportunity that arises, no matter how far fetched (such as the vocal organs). Thus in a Platonic sense language is the shadow of thought. Or is thought not possible without language? Is the biological emergence of language a pre-requisite for consciousness? Such considerations do not enter into Crystals book. In short, it does not really address the issue of how language works.