economical and elegant. In this case, unlike the one of reverse languages, we humans certainly do not "lack" the intelligence needed to acquire and use such constructions. After all, with the aid of paper and pencil and with all the leisure of time and memory, we can reconstruct the meaning of such sentences, just as we can learn BASIC, FORTRAN and LISP and write programs vastly more elaborate than the above simple sentence. But, plainly, we are constitutively unable to access these sentences, presumably because of language-parsing problems. Actual communication based on constructions such as the above, no matter how expedient and quick, is precluded for us by some quirk in our perceptual and computational makeup. The actual principles of universal grammar would license center embedding, yet we are unable to profit from this liberality because of other constraints acting on memory, parsing, perception and computation.

Once more, there is no way of deducing any of the above facts (and plenty more which I cannot review here) from adaptive criteria, either alone, or conjoined with criteria of simplicity, expediency, minimization of memory-load or of computational burden.

We have to conclude that popular notions such as "general intelligence," "communicative function" and "problem-solving" are totally useless in understanding the origins and nature of human language. The Darwinian adaptationist picture, in fact, not only fails to explain even the most central features of language and mind, but it forces upon us intuitions that are highly misleading. When, on the contrary, we adopt an exaptationist picture, these constrictions and fallacies have no reason to apply and we can at last reconcile what has been independently discovered in the domain of language and mind with a biologically credible story.

It is out of question that I should even attempt to summarize here what has been independently discovered. However, since the main target of this paper is the instructivist conception of learning, even a few simple linguistic facts may suffice to support my argument.

8. Why instructivism could not work

The following is a rough intuitive sketch of the reasons why the instructivist picture of learning is overwhelmingly likely to be wrong in the realm of language and mind just as it was proven wrong in biology. It will be, I hope, self-evident that it immediately generates a number of difficulties, some of which are insurmountable. The conclusion will be that language "learning" is something that "happens" to the child and not something that the child "does" (Chomsky, 1986b, 1988).

What follows is certainly very far from a "demonstration" of this thesis, nor do I pretend to provide a knock-down argument. I will just pick out the most commonly invoked instructivist mechanisms of learning and show, through very simple linguistic examples, how each fails to account for the richness, the depth and the extreme specificity of one essential component of the language faculty. But this is just for the sake of brevity and exposition: The whole sad truth for the instructivist is that each turns out to fail badly for all components.

Imitation: The counter-case of phonology

Unless one grants that the child starts out with a strong, specific, a priori (i.e., innate) predisposition to select who is to be imitated, in what respect, to what extent and when, one is immediately led to a host of paradoxes and wrong predictions. It is well known, for instance, that children of immigrants grow up totally devoid of accent, no matter how heavy the accent of their parents may be, when they are raised in a homogeneous community of native speakers. Why don't they imitate the accent of their parents? Why do bilingual children never try out an unseemly mishmash of accents and sounds extracted from both languages at the same time? Why do they "switch" so neatly from one language to the other?18

The literal role of imitation in phonology is, moreover, seriously disqualified by the very nature of speech-sounds (Chomsky & Halle, 1968; Halle, 1985, 1987, 1988; Halle & Vergnaud, 1987; Liberman, 1982), and by a host of recent findings in the domain of neonate cognition (Mehler, 1985, 1987; Mehler & Fox, 1985; Mehler, Lambertz, Jusczyk, & Amiel-Tison, 1986). The newborn (and even the prematurely born) infant displays a highly sophisticated array of specific linguistic filters allowing for instant discrimination between linguistic and non-linguistic sounds. As recently shown by Jacques Mehler (Mehler et al., 1988), 4 days after birth the neonate already discriminates between the voice of the mother and the voice of another woman of the same age, between a natural flow of discourse and a chopped sequel of isolated words, even between the language of the mother and another language (sentences of French versus Russian, or of Italian versus English, as

18 They do occasionally borrow from the lexicon of the other language, but these are either direct "quotations" due to lack of knowledge or memory, or shotgun phonetic adaptations of foreign words that are then submitted thoroughly to the phonological rules of the target language. The plural of, say, "pizza" is "pizzas," but of "blitz" is "blitzes," in agreement with English rules. As Morris Halle (1985, 1988) has rightly emphasized, these shotgun adaptations do not arise from any imitation, but ensue automatically from the largely innate knowledge of quite abstract and intricate phonological constraints proper to the target language.
normally uttered by one and the same perfectly bilingual person). At 7 months of age the infant already perceives clauses as distinct units of particular relevance in the natural speech flow (Hirsch-Pasek et al., 1987).

Full-blown phonological analysis of the restrictions that each language severely imposes on possible sound sequences in that language (that is, on admissible syllable structures for that language), on stress patterns, on tonal profiles, morphology and much besides (Halle, 1988), shows that all known human languages are subsumed under very abstract and universal principles of "parameter fixation." There is a fixed number of such phonological parameters and each has only two admissible positions. Poor and occasional input from the surrounding community of native speakers is sufficient to tell the child how these parameters are to be fixed. The very rich, specific, innate phonological predispositions of the child allow him/her to fix these parameters correctly, upon minimal episodic exposure to the sound patterns of the community.

Another significant recent discovery of phonology (Halle, 1988) is that the fixation of these highly abstract parameters maps directly onto the gymnastics of the physiological articulators of the vocal tract (the tongue, lips, soft palate etc.) and that, therefore, the mental representations through which words are memorized and the unconscious instructions being imparted to the articulators share a common format. In other words, we may represent the situation at an intuitive level as if there were sets of binary choices (+ or -) to be tacitly, that is unconsciously, assigned by each native speaker of each natural language (or dialect) to the value of each of a number of phonological parameters bearing on pronunciation, on syllable structure, intonation, stress, as well as on standard morphological features such as plurals, passives etc. These sets of binary choices specify very abstract properties of the speech sounds proper to the natural language in question. It turns out that they also, almost by the same token, specify the admissible degrees of freedom in that language of the physical movements of the various articulators. Since this close parallelism between abstract phonological...
so forth. Nor do they err in detecting a foreign or "impossible" term, when they hear one. They also stick to the lexical morphology of their language and make few mistakes in assigning lexical categories. They tend to generate idiosyncratic lexical constructs ("I frowned to school," "Pour the glass with lemonade," "I wipe the table dirty" etc.) (Bowerman, 1982; Gropen & Pinker, 1986; Pinker, 1986, 1989; Pinker, Lebeaux, & Frost, 1987), but they show an amazing competence in generating quite "plausible" idioms.

Let's come to what might, at least prima facie, be explained by association and induction. Here is the standard situation of lexical acquisition: The child is confronted with a certain state of affairs (a real-life situation, a picture book, a movie, a tale etc.) and hears a new word, usually embedded in a sentence. There are too many possible candidates around, in the situation, for the child to pick out exactly what the new word "stands for." He does not know with what to "associate" it. (This is, incidentally, another telling instance of the "poverty of the stimulus." So he or she asks. And the adult severally manages to provide the meaning:

**First striking fact:** by hearing just one sentence, and only once, the child already picks plenty of cues as to the possible (and, very interestingly, also the impossible) meanings of an unknown term. 21 Within the remaining degrees of uncertainty, there are innumerable alternative right ways to convey the exact meaning of a given word, but the child will "get" it, irrespective of the particular right way one chooses.

**Second:** the child will thereafter "possess" that meaning, generalizing often without fault well beyond the initial situation and beyond the contingent introductory device adopted by the adult (Armstrong, Gleitman, & Gleitman, 1983; Landau & Gleitman, 1985).

**Third:** for a number of years, up to age five or six, the child acquires some 10 new words every day, one for every waking hour, day after day, month after month (for a recent assessment, see Miller, 1986; Miller & Gilda, 1987).

**Fourth:** by just getting the lexical meaning right, the child automatically knows a lot of other things about that meaning.

**Fifth:** the child automatically rules out countless meaning-candidates that are perfectly thinkable, easily expressed through short sentences or circumlocutions, that represent perfectly ordinary situations of everyday life, but which **cannot** be the meaning of any single lexical entry in any natural language (for instance: No single verb entry can mean "to eat bread and..., "to read a book, but not..., "to eat all the bread first, and then...”; Hale & Keyser, 1987; Pinker, 1987; Pinker et al., 1987; Tenny, 1987).

Many straightforward examples (see notes 19, 20, 21) illustrate these core facts about the acquisition of the lexicon. Very little, if anything at all, here is "learned" and "taught." Many of these core facts about the actual, the possible and (most important, from our present point of view) the impossible lexical meanings apply uniformly to all human languages. But also a lot of the language-specific facts (i.e., those unique to English, Italian etc.) are equally untaught and they just come along, mandatory and at no price, once the lexical meaning is acquired. 22

Association and induction cannot even begin to account for this kind of knowledge, which is acquired invariably, without effort and with quite marginal errors of under-differentiation and over-generalization (leaving aside the trivial phenomenon of occasional mismatch between a certain label and its lexical meaning).

The most plausible explanation 23 is that these subtle (and probably highly inter-connected) facts about lexical meanings are known **innately**. The process of "learning" just picks them out en bloc, once the phonetic label is assigned to the concept. Association now becomes a very minor component of the mechanism: What is being associated is just the sound for that concept and the innately available concept itself. 24

It is no mean feat, I think, to have come to the conclusion that we must construct an innatist-selectivist theory even for the acquisition of the lexicon. Association is marginal to the process; it boils down to a heavily constrained pairing of certain labels to certain innate concepts.

---

21Let's take an imaginary verb (because this helps us reconstruct the process) embedded in a standard sentence: "Daddy splonked the fish in the kitchen at noon." Just from this sentence the child deduces plenty of facts about the meaning of the unknown verb: that it is transitive, that it expresses something that can be done to fish, that daddy did it at noon, that this something can be done in the kitchen, in a brief period of time, that it does not require more than one person to do it, and so forth. These simple facts already automatically and **obviously** rule out an infinity of meaning-candidates for this verb (Higginbotham, 1986).


23The above sketches should suffice to suggest how heavily and jointly constrained is the association process by the innate phonological, syntactic and semantic structures and by the values assigned to the parameters in the local language.
Analogy, generalization and problem-solving: The counter-case of syntax

There is, as we will see, a very strong case against any process of language “learning” as based on surface analogy (that is, on analogies between the manifest, acoustical shape of utterances). The strong case is based on the impossibility to utilize even surface identity, which obviously ought to be a clear-cut, indisputable and more “objective” sub-class of such analogies. Yet, I will start here with a more intuitive case against analogy: one based on the inevitable sequel of disillusionments which the child ought to face, if she were indeed to use analogy and “problem-solving” in the acquisition of her native language.

Early on the child “learns” the relation between the meaning of the following pairs of very common sentences:

(1) John ate an apple
(1a) John ate (understood: ate something or other)

(2) John received a gift from his mother
(2a) John received a gift (understood: from someone or other)

(3) John cuts the bread
(3a) John cuts (understood: generic activity of cutting)

Let’s try a tentative simple-minded rule, derived from analogy: When you delete the object in a sentence containing a transitive verb, the resulting sentence means that the action described by the verb is expressed in a generic way, object unspecified. Therefore the child should naturally, by analogy, from

(4) John wears a hat
(4a) *John wears

(which is supposed to be understood: wears something or other). But (4a) is not a well-formed sentence in English.25 In fact we never even try to produce (4a), nor does the child. Why? Object deletion is likewise forbidden for “say” and “tell,” but not for “speak” or “write.” The incorrect constructions, formally equivalent to (4a), are not even tried out. Again, why?

25This section contains variations on, and drastic simplifications of, topics thoroughly discussed by Chomsky in his book Knowledge of language (Chomsky, 1986a).
26The same applies to its translations into many other languages (Kenneth Hale, personal communication, September 1987). It applies, for sure, to French and Italian.

The next step might be, according to the analogy and problem-solving hypothesis, to restrict the class of verbs for which object deletion is permitted.27 But, certainly “catch” and “wash” do permit it, and in fact (5), (5a), (6), (6a) are all well-formed sentences:

(5) John catches rabbits
(5a) John catches (something or other)

(6) John washes his hands
(6a) John washes (generic activity)

Yet, things go wrong when we try to generalize (assuming that this is what we try to do) to other sentences containing these verbs:

(7) John is too clever for us to catch Fido (Fido is, say, John’s dog)
(7a) John is too clever for us to catch

Plaintly, analogy collapses: Any speaker of English immediately understands that (7a) does not mean that John is too clever for us to catch something (or someone) or other, but that he is too clever for us to catch him. John. We can try out an analogy-amendment: Object deletion in subordinate clauses introduced by “for X” has the effect of discharging the action of the verb onto the subject of the main sentence.28 But this does not work either:

(8) The pool is too cold for us to wash the car
(8a) The pool is too cold for us to wash

Now, (8a) does not mean that the pool is too cold for us to wash something or other (and the amendment now grants that this is not the case), but (8a) does not mean, either, that it is too cold for us to wash it, the pool. The procedure by analogy goes, once more, badly wrong. Every speaker of English understands (8a) to mean that the pool is too cold for us to wash our—

27Maybe it has to do with verbs that express actions that modify the state or the material integrity of the object (eat, drink, write, shave, wash) and verbs that don’t (wear, tell, say). The point is that this highly elaborate knowledge of subtle lexical facts (whatever they are) has very crucial consequences for syntax and semantics in general; these facts are naturally accessible to all of us, quite unconsciously and without any explicit “teaching” (Hale & Keyser, 1987).
28Consider, however, that this cumbersome explicit (or nearly explicit) procedure of trial and error is exactly what the child is supposed to be doing when she learns the language (under the hypothesis of an analogy-driven mechanism). Indeed, the current state-of-the-art in generative grammar attributes to the subject computations of a degree of abstractness and intricacy that dwarf the above, by orders of magnitude. The point is, however, that these are supposed to be mandatory, totally unconscious and innately available to us all. The flaw of an explanation of language learning based on “analogy” and problem-solving lies not in the complexity, but in the conscious corrigibility, in the “transparency” to revisions of the mental procedures that it has to attribute to the subject.
selves. Object deletion has produced yet another totally different effect. This effect is perfectly well known to all speakers: The child does not, as a matter of fact, interpret (7a) and (8a) wrongly, under the spell of the alleged analogy, and he/she does not need explicit corrections.

In a desperate move, the lover of analogies might try to sharpen the tentative rule in the following way: Object deletion in subordinate phrases introduced by “for X” has the effect of discharging the action of the verb either onto the subject of the main sentence, or onto “X.” Admittedly, the rule is a bit sloppy, but we may work hard to refine it; this is how analogy and problem-solving works.

But, let’s now try out other, slightly more complex sentences:

(9) Lisa considers John too clever to expect us to catch Bill
(9a) Lisa considers John too clever to expect us to catch

If we accept (9a) as a well-formed sentence at all, then it restores the original effect of object deletion. Catch, once again, means catch somebody (or something) or other. To realize that this is the case, consider (10)

(10) Lisa considers John too clever to expect us to shoot

Where “shoot,” plainly, stands for the action of shooting (target unspecified). We are back to square one. Object deletion functions here as in the initial, very elementary cases.

What must be baffling to those who believe in analogies and problem-solving is that we have no uncertainty whatsoever in understanding these sentences the way we do. Nothing prompts us to postulate that the child carries out a “search and find” procedure, figuring out all sorts of tentative meanings through analogy, and then revising provisional hypotheses under the impact of corrections received from other speakers.

Chomsky’s argumentative strategy in his case against analogy (1986a, 1986b, 1988) is even more drastic than the above. He denies the relevance of similarity and analogy in the acquisition of language by showing that not even identity is preserved. For instance, both (7) and (7a) contain the string “John is too clever for us to catch,” repeated without variation. Yet, this string is differently interpreted in (7) and in (7a), because the underlying structures are, in fact, different. Other striking cases offered by Chomsky (1986b, 1988) are:

(11) John expects to feed himself
(11a) I wonder who John expects to feed himself

It is quite plain to every speaker of English that, while “himself” refers to John in (11), it just cannot refer to John in (11a). Yet there is superficial identity between the string “John expects to feed himself” in (11) and in (11a). Since identity is a narrower, more stringent, and more “objective” case of analogy, and not even identity is preserved across linguistic inputs, then analogy can be of no use to the language “learner.”

To complete this (admittedly) summary reductio ad absurdum of analogy and problem-solving, we have to dispose of another recurrent temptation: the canonical appeal to “expectations,” to “information content” and to our “knowledge of the world.” Semantics is not governed by statistical expectations. We experience no more difficulty in understanding the highly unexpected sentence (12)

(12) I have a giraffe in my boot

than in understanding the highly expected sentence (13)

(13) I am coming to the conclusion.

Similarly, what we know about the world jibes well with the meaning of

(14) Most of those who own a donkey beat it, but nobody who owns a parrot does

while it does not jibe at all with the meaning of

(15) Most of those who own a parrot beat it, but nobody who owns a donkey does

But understanding (15) does not pose any special problem. The meaning of (15) is a perfectly fine meaning, though it does not express a fact that we know to be true of the world. The true and very intriguing linguistic fact is that (14) cannot express what (15) expresses; why? These are basic facts about ourselves and our knowledge of language which, plainly, cannot be explained in terms of expectations, general intelligence, information and generic knowledge of the way the world is.

When one considers legions of cases like these one cannot fail to see, at least in my opinion, that all attempts to explain the learning of language through analogy, problem-solving and induction prove utterly implausible. Once again, we are confronted with a process of selection, not one of instruction.
9. A summary picture

In outline, what we get is the following picture. Our species innately possesses a rich, specific, modular and highly articulate capacity for language, organized around certain universal “principles.” These principles allow for a well-determined variation, through the choice of discrete values for certain parameters. In order to acquire a specific natural language (Chinese, English, Russian etc. or, rather, some specific local subset—a “dialect”—of any of these) the principles of UG have to be severely “parametrized” through the linguistic input from the community to the child. In the absence of disconfirming evidence, the parameters are innately set to their “unmarked” values (Flynn, 1987; texts in Roeper & Williams, 1987; Wexler & Borer, 1986; Wexler & Manzini, 1987). For certain languages, certain unmarked values are stabilized indefinitely, throughout the entire life of monolingual native speakers. Other parameters, instead, will eventually switch to the new “marked” value and then remain “set.” Without any loss of generality, we can idealize this switching process as taking place instantaneously, upon one single exposure to the relevant linguistic input.

The principles of UG define what is and what is not a “possible” human language. These principles are totally distinct from the (equally innate) principles governing gesturing (Petitto, 1986, 1987), vision (Marr, 1982; Ramachandran, 1987), mental imagery (Kosslyn, 1980; Shepard & Cooper, 1982), movement (Cheng, 1986; Cheng & Gallistel, 1984), concept acquisition (Carey, 1985; Keil, 1979, 1986; Smith & Medin, 1981; Spelke, 1985, 1986), counting (Gelman & Gallistel, 1978; McCloskey, Sokol, Goodman-Schulman, & Caramazza, 1987) and other mental faculties (Fodor, 1983; papers in Garfield, 1987). As we saw, there is no logical necessity to guide us in understanding these linguistic principles, and it is easy to conceive systems of communication that might have equally well secured our survival adopting very different principles.

We can safely grant that acquiring systems of pragmatic rules, a general knowledge of the world, an understanding of other people’s beliefs and intentions and of social roles may be directly adaptive (may increase our Darwinian fitness) but the disparaging fact, for the adaptationist, is that each one of these aptitudes presupposes in theory, and exploits in practice, all the other aptitudes. It is very unclear what created “selective pressures” for what.

It is minimally fair to admit that at present nobody knows what is involved in learning algebra, quantum mechanics or American history. There seems to be, here, a prima facie case for some learning, perhaps without quotation marks. All we can say is that these cases of acquisition, requiring years of intensive application and hard conscious efforts, are drastically different from our spontaneous, effortless acquisition of language. I hope that I made clear why. My aim here was to draw some positive lessons from what we know, rather than extrapolating defeatism from what we don’t.

Conclusion

The demise of learning undoubtedly sounds like a paradox. Yet, this apparently paradoxical state of affairs is no more paradoxical than it once was in evolution, in immunology, in enzyme action and in neurobiology.

All I wanted to suggest is that, since the tacit assumptions still dominant in linguistics and psychology the world over are an exact replica of those once shared by most biologists against selective theories in biology proper, then a conceptual shift in the same direction can be expected to take place in linguistics and psychology too. And for much the same reasons. Paradoxes are often explained away when we denounce some tacit and “innocent” assumption. My argument here was not one for biological reductionism, but rather one against unwarranted innocence.

References

Evolution, selection and cognition

Evolution, selection and cognition


M. Piattelli-Palmarini


**Résumé**

La plupart des biologistes et certains cognitivistes sont arrivés indépendamment à la conclusion que l'apprentissage, au sens instructiviste et traditionnel du terme, n'existait pas. Cette thèse peut sembler extrémiste, mais je la défend ici à la lumière de données et de théories provenant d'une part de la biologie, en particulier de la théorie de l'évolution et de l'immunologie, et d'autre part, de la grammaire générative moderne. Je souligne également que la chute de l'apprentissage est incontestée dans les sciences biologiques, alors qu'un consensus similaire n'a pas encore été atteint en psychologie ni en linguistique. Puisque de nombreux arguments offerts à l'heure actuelle en faveur de l'apprentissage et d'une capacité d'"intelligence générale" s'appuient souvent sur une image déformée de l'évolution humaine, je dévoile quelques sections de cet article à une critique de l'"adaptationisme", en donnant également les éléments d'une meilleure théorie de l'évolution (fondée sur l'"exaptation"). De plus, certains arguments en psychologie et en intelligence artificielle présentés aujourd'hui comme indubitalement sont en fait une réplique exacte des anciens arguments en faveur de l'instruction et contre la sélection en biologie; je m'appuie sur ces erreurs du passé en tirant des leçons pour le présent et le futur.