Comments on papers by McNeill and Foppa

The origins of language and language awareness

WILLEM J. M. LEVELT

Both papers in this final section are about the origins of language. It was with great hesitation that I agreed to be a referee, since discussion about the origins of language, which are as old as our culture, have mostly been absolutely futile and a sheer waste of time. In order to protect its own members, the Linguistic Society of Paris prohibited any communications about the origins of language in 1866. However, this did not help for long. Darwin's *The descent of man* was published in 1871, and it produced a new era of discussion, in which we are still involved.

This century of discussion has, again, taught us very little about the phylogeny of language. The main finding, which was certainly unexpected in the Darwinian framework, was that all primitive peoples turned out to have full-fledged vocal languages. Both the full-fledged and the vocal were surprising. If language evolution did develop gradually by natural selection, no trace of this process can now be found. A one-step mutation, as implied by Chomsky (1968), is biologically untenable in view of the complexity of the behavioral patterns involved. The vocal modality of all human language was a surprise because the general feeling after Darwin was that language must have been gestural in origin. Many philosophers, linguists and biologists have argued this long before McNeill's contribution to this volume. I mention, for instance, Bulwer, who in 1644 called gesturing the 'onlely speech that is naturall to Man', also Rousseau (1782), Wundt (1900), and there are many others. Again, although gesture is everywhere important in language communication, no tribe was ever found where language was essentially gestural,[1] in spite of expectations to the contrary (see 'observations' made by Kingsley (1897) and dismissed by Langer (1942)). There are, moreover, no other anthropological sources of evidence for a gradual transition from gestural to vocal language in the human species. And the arguments by Hewes to which McNeill refers are without force. Tylor concluded in 1878 that 'The idea that the gesture-language represents a distinct separate stage of human utterance, through which man passed before he came to speak, has no support from facts.' This conclusion is still valid.

So McNeill cannot use evolutionary findings as an argument for his gestural theory of language origins. Rather, McNeill's phylogenetic claims are no more than hypotheses which are exclusively based on evidence from ontogenesis and from experimental psychology. The evidence is interesting, but inconclusive, as I will argue in a moment. The main point, however, is the futility of such discussions: what can be gained by inferences from ontogeny to phylogeny, as long as anthropology leaves us, a century after Darwin, fully in the dark about the evolution of human language? Since claims about the phylogeny of human language can at best be paraphrases of other empirical findings, the only gain is a false suggestion of additional support to such findings. And this happens in McNeill's paper where he cites Goldin-Meadow's work. She found spontaneous development of gestural language in deaf children. McNeill concludes that this 'is consistent with the theory that vocal speech overlies an ability for manual communication. When vocal speech is blocked the original gestural system remains.' It is, however, equally consistent with other theories, notably the one which claims that if the vocal channel is blocked, the most flexible alternative will be used for linguistic communication. In deaf-blinds, for instance, gesture language is also impossible, and a manual-tactual replacement is developed, but it seems far-fetched to conclude from this that this must be the evolutionary more primitive form of linguistic communication. Since I see no evidence whatsoever for McNeill's gesture theory, I need not ask him why he didn't deal with the main problem of this theory, namely, if gesture is such a natural system, why did the human species switch at all to the horrors of a vocal language?

---

[1] It is true that gesture languages have been found, but vocal language is always the main mode (e.g. Plains Indians, Amazon Indians, Bushmen, see Stokoe 1974).

So much for the gestural origins of language. Let us now move to the essence of McNeill's paper, his syntagma theory. It says, if I understand it well, that the ontogenesis of language is based on the unique human capacity to link the articulation system to the system of sensory motor representations. In my view McNeill touches an essential point here. The fact that humans, and also chimpanzees for that matter (e.g. Premack 1977), can communicate linguistically about things, states, actions, or events not present is in full agreement with present-day cognitive psychology, in that one has to suppose elaborate systems of internal representation of such things, states, actions, or events. Major experimental evidence for the structure of such internal representations, and for operations within such systems, comes from various sources, among them from reaction time studies. By the way, this method is – to my knowledge – not frequently practiced in ethology. This is regrettable in view of the high precision by which otherwise elusive behavior patterns can be measured.

McNeill stresses the enormous adaptive value of using such systems, i.e. of being able to perform 'vicarious actions'. This is the ability to do internal or covert computations on internal representations, instead of having to perform overt actions to check out the consequences of one's behavior. If a linguistic signalling system of vocal or gestural signs is attached to such internal representations and operations, this would allow for very easy transmission of action plans, directions, etc. within a community. But again, McNeill comes close to circularity in his exposé, this time because of the use of rather ill-defined terms. I would like to mention one example of this. McNeill argues that the syntagma theory would predict that output integration should coincide with meaning integration, and he does not hesitate to add that 'this result has been found'. There is much here that the reader has to take on trust. On the one hand McNeill leaves undefined what is meant by 'a single meaning structure', and how it could be manipulated independently. On the other hand, McNeill doesn't tell us what units of speech articulation are. Are they related to Ekman's 'punctuation' (see chapter 3.1 above), or to what Goffman called 'topical runs' (chapter 3.2 above)? According to McNeill, they are not the familiar linguistic units such as sentence or clause, and one is left with the impression that they are really those stretches of speech that correspond to single meaning units. But this would make the argument fully circular. Since Professor McNeill cannot have meant that, I would like to ask him to define his constructs in such a way that I can understand what is meant by 'this result has been found'. It might especially be helpful if he could give reference to empirical work on which his statements are based.

In McNeill's paper the articulation system is sketched as a late and somewhat arbitrary addition to the system of sensory motor representations, and it is left undiscussed. Foppa's contribution is a good complement to McNeill's in that it centers around the ontogenesis of articulated speech. Foppa starts out by sketching how little ethology can tell us about the evolution of human language, especially since it is so very different from spontaneously developing communication systems in animals. So, Foppa has also to restrict himself to ontogenetic evidence. The ethological question as to the role of natural selection shrinks to the question as to whether there is a mechanism of Skinnerian shaping in the child's language acquisition. Again, the answer to the latter question will, in my opinion, tell us nothing about the former, in spite of possible analogies. Foppa is wise enough not to try this route seriously.

The main issue in Foppa's paper is the issue of self-correction in children's speech. Here, Foppa and his co-workers have developed a method which can metaphorically be called 'ethological'. The adult says 'mhm?', suggesting that he doesn't understand the child, and the child's repeat performance is registered. As usual, the method is not really new. I have seen similar ideas in a 1914 paper by Bohn, and similar experiments in recent papers by Grimm (1975), Stokes (1976, 1977), Cherry (in press), and Garvey (in press).

I think this type of work is very important: it stresses a rather typical human aspect of language acquisition, and it might be of inspiration to anthropology, and even to ethology. I will finish my introduction by sketching a more general framework for this research, and my question to Professor Foppa is whether he agrees to be placed in this pigeon hole.

The framework is called 'metacognition'. One talks about metacognition if there is evidence that a person has knowledge about his own cognitive processes. A typical case is metamemory: the person knows that he is going to forget and therefore he makes notes, or constructs a mnemonic. Meta-activities are very general in humans, and the issue is not really different from what philosophers have called 'self-consciousness'. Especially Flavell (1976, 1977) and Brown (1977) have done empirical work here, and this has led them to distinguish different levels of metacognition. The lowest level, which is apparent early in ontogenesis, is *self-monitoring*: making self-corrections, either spontaneously, or as a response to specific stimulation. The highest level is explicit *reflection* on one's own cognitive activities and products, such as explaining how one solved a problem.

Metalinguistic behavior is a special case of metacognition. Foppa mentions one higher level of this which he calls 'metacommunication', but limits his study to the lowest level: self-monitoring in the child's speech. The important question to ask about metacognition is: what function does it fulfill in the acquisition of cognitive skills, i.e. in language development? Skinner has little to say about this, and Foppa suggests that it might help to establish internal standards in a process similar to concept learning. One function of linguistic awareness in the child might be to shape his internal representations of linguistic tools by vicarious action, to use McNeill's terminology. The case described by Weir, and cited in both papers, is an example of this shaping of linguistic tools. Nothing much is known about the growth and functions of linguistic awareness in the child, but I can refer to a book which we recently edited in our Max-Planck project-group (Sinclair, Jarvella and Levelt 1978).

Metacognition is probably universal in humans, but detailed anthropological work is needed to determine its bounds and cultural variations. A first anthropological paper in metalinguistics has been written by Heeschen (1978); it analyzes the linguistic means by which the stone-age Eipo can refer to their own language.

Back to Foppa's specific case: the child says something to the adult, and the adult feigns not to understand. The child is aware of the failure of his utterance and makes a new one, expressing the same intention.

It seems that, psychologically speaking, linguistic awareness arises out of devices for finding faults. These cognitive devices could do two things: (i) detect that a fault has occurred, and (ii) specify what sort of fault it was. An intriguing analysis of such devices can be found in a paper by Marshall and Morton (1978), from which most of the following notions and examples are derived.

Fault-finding devices operate both in the production and the perception of language. Foppa's examples concern *production*. Normally, the child would formulate his intended message and receive feedback that he has been understood. No specific awareness needs to be involved. If the feedback (or absence of it) signals that the utterance has failed, the child will send the same message into his formulating system again, and a new utterance arises. This new utterance may differ from the original one either through random variability in the formulating system, or through systematic attempts by the child to change the structure of his formulator. Foppa's cases are examples of fault-detection, but fault-specification can also occur in language production, for instance in spontaneous self-correction, such as: (child, age 2; 2): 'Look, look. A caterpillar . . . helicopter (laugh).' Such self-corrections are frequent in early spech and need explanation. In the example, the child becomes aware that the meaning of the word his formulator produced (caterpillar) is different from the meaning he intended to express (that of helicopter). One has to assume that the child monitors its speech by extracting the real meaning of caterpillar after it has been produced, then compares this meaning to the intended meaning and finds the fault. This involves, therefore, a feedback loop and a comparison device.

Very clear cases of such early linguistic awareness can be seen in the child's *perception* of language. One example (again from Marshall and Morton) should suffice. The child (age 4; 7) comes home from school, and enters the very noisy kitchen.

Adult: Did you enjoy yourself?
Child: What?

At this point the child has detected a fault. There was speech input, but no meaning-output. The continuation of the conversation shows that the child has moreover identified the sort of fault:

Adult: Did you have a good time at. . .?
Child: No, no, *say* it again!

The adult guessed that the problem was one of lexical understanding, but the child's answer makes it clear that his problems were merely acoustic. The child was thus able to locate the source of the communication failure, showing awareness of stages in the process of speech perception.

What sort of functions could be fulfilled by such metalinguistic devices? First, they seem to improve the *ad hoc* communication, which is certainly advantageous. But secondly, they may also have a specific function in language learning. The mentioned feedback loop from perceptual parser to formulator may be an important channel for the parser to 'teach' the formulator, so that the perceptual detection of a fault leads to a permanent change in the formulator. The young child's capability of language perception is far ahead of his production competence. Linguistic awareness might help him bridge this gap.

Similar mechanisms of awareness may be involved in the acquisition of other cognitive skills as well, and an interesting question for ethology is whether such phenomena can also be observed in animals. The closest may be the acquisition of skills through play – play is an outstanding case of becoming aware of a tool, be it a physical or a cognitive tool. But could one observe the use of meta-devices in the acquisition of a communica-