

Can there be a Language of Thought?*

(In: R. Casati, B. Smith and G. White (eds.) *Philosophy and the Cognitive Sciences. Proceedings of the 16th International Wittgenstein Symposium*. Wien: Hölder-Pichler-Tempsky 1994, 207-219)

Ansgar Beckermann

1. Cognitive sciences in a broad sense are simply all those sciences which concern themselves with the analysis and explanation of cognitive capacities and achievements. If one speaks of *cognitive science* in the singular, however, usually something more is meant. Cognitive science is not only characterized by a specific object of research, but also through a particular kind of explanatory paradigm, i.e. the information processing paradigm. Stillings *et. al.* for example begin their book *Cognitive Science* as follows:

Cognitive scientists view the human mind as a complex system that receives, stores, retrieves, transforms, and transmits information. (Stillings 1987: 1)

The information processing paradigm however, leads directly to the paradigm of symbol processing, because a system can, as it seems, only receive, store and process information if it has at its disposal a system of internal representations or *symbols*, i.e. an internal language in which this information is encoded. At least this appears to be an idea which suggests itself and which Peter Hacker expresses as follows:

... if information is received, encoded, decoded, interpreted and provides grounds for making plans, then there must be a language or system of representation in which this is all done. (Hacker 1987: 486f.)

And indeed the assumption that in cognitive systems there must be something like a system of internal representations, or a language of thought,¹ lies at the heart of many new works in the fields of cognitive psychology and cognitive

* This is a revised version of a German paper which I read at the „Jahrestagung des Instituts für deutsche Sprache“ on March 16, 1993, in Mannheim. I would like to thank Antonia Barke for translating the paper into English.

¹ The expression ‘language of thought’ (*lingua mentis* - ‘*Sprache des Geistes*’) which was – as far as I know – first used in this context by Harman 1973 is seriously misleading because the expressions of the language of thought are – and Fodor e.g. agrees with this – internal *physical* states of the respective system, as for example certain neuronal firing patterns or bit patterns in the memory of a computer. Hence expressions like ‘language of the brain’ or ‘language of the computer’ would be more precise.

neurobiology. For these sciences this assumption has the status of an empirical hypothesis, that is to say, for them, internal representations or symbols are theoretical constructs which are postulated because they allow us to explain cognitive achievements in a well corroborated and systematically particularly satisfying way. On the other hand, there are philosophical approaches that support the assumption through very general considerations concerning the nature of mental states.²

These and other related approaches have been criticized by a variety of authors in many different ways. Especially in Oxford however, criticisms have been formulated which are based on the late Wittgenstein and which radically question the symbol processing paradigm in general.³ Peter Hacker, for example, in his article „Languages, Minds and Brains“ asks the rhetorical question:

Is this [sc. the idea that there is a language in the brain] just a picturesque metaphor or helpful analogy? Or is it a symptom of widespread confusion in the presentation, description and explanation of experimental data ...? (Hacker 1987: 487)

And his answer indeed states, that the idea of a system of symbols in the brain is founded upon a fundamental confusion of concepts and therefore is literally *nonsensical*. What are Hacker's reasons for this devastating assessment?

His argumentation begins with a characterization of the idea he then wishes to attack:

The general conception at work involves the supposition that the brain has a *language* of its own, which consists of *symbols* that *represent* things. It uses the *vocabulary* of this language to *encode information* and it produces *descriptions* of what is seen ... (Hacker 1987: 488)

A 'symbolic description' is presumably an array of symbols which are so combined as to yield a true (or false) characterization of a certain aspect of the world. It must be cast in a certain language which has a vocabulary and grammar. (Hacker 1987: 488)

We, thus, have to ask what it *could* mean for the brain to possess a language with its own vocabulary and its own grammar. Before trying to answer this question however, we should first get clear about what it *does* mean in general to say that someone possesses a language.

² The main figure in this field is certainly Jerry Fodor, who developed his Representational Theory of Mind over many years before casting it into its canonical form in *Psychosemantics*. See Fodor 1975; 1978; 1981; 1987.

³ See Hacker 1987, and also the new collection of essays Hyman 1991.

Someone who *has* a language has mastered a technique, acquired or possesses a skill of using symbols in accord with rules for their correct use, or – if you prefer – in accord with their meaning. (Hacker 1987: 491f.)

Someone's having a language thus consists in his possessing certain abilities. He understands utterances made in the language; he knows the meaning of the words of this language and is able to use them in order to carry out a broad variety of speech acts: He can call a taxi, ask for the way to the rail station, tell stories or make jokes, order wine with a meal, introduce a friend, describe a landscape, and so on. Over and above all that he can – should it happen that he is not understood – explain what the words he has used mean and what he wanted to say by uttering them.

If [someone] understands a language he can respond in various ways to others' uses of words and sentences, as well as correcting others' errors, querying their unclarities and equivocations. (Hacker 1987: 492)

From this fact alone – that mastering a language implies all these skills – it follows, according to Hacker, that it is literally nonsensical to say that the brain possesses a language.

Only of a creature that can perform acts of speech does it make sense to say that it has, understands, uses, a language. But it is literally unintelligible to suggest that a brain, let alone *a part of a brain*, might ask a question, have or express an intention, make a decision, describe a sunset, undertake an obligation, explain what it means, insist, assert, instruct, demand, opine, classify, and so forth. (Hacker 1987: 492)

In order to be capable of possessing a language one must be able to carry out certain actions – actions which belong to a different level than those from which one can meaningfully say that they are being done by a brain, let alone parts of a brain. Brains or parts of brains are not therefore, for conceptual reasons alone, possible language users.

But there are more reasons which, in Hacker's view, show that the idea of a language of the brain becomes increasingly absurd the more the implications of this idea become clear to us. The expressions of a language, he continues, have a use governed by convention and someone who masters a language must know the correct use of these expressions, i.e. he must be able to distinguish correct uses from incorrect ones.

A rule-guided use of language which refers to standards of correctness can, however, only be founded on a social practice.

For only where there is a practice of employing a sign can there also be an activity of matching the application of the sign against a standard of correctness. Since signs have a meaning, a use, only insofar as there is a convention, a standard of correctness for their application, there must be a *possibility* of correcting misuses by reference to the standard of correctness for the use of the expression which is embodied in an

explanation of meaning. The use of language is essentially a normative activity. (Hacker 1987: 496)

This is another reason why according to Hacker it is altogether impossible that brains or brain cells employ a language: One cannot meaningfully say that brains or brain cells follow conventions, because conventions can only be followed if they exist at all. They can only exist however, where they are used within a social community in order to teach and learn, to correct mistakes and explain and justify actions.

Only of a creature who has the *ability* to make a mistake, who can *recognize* his mistake by reference to a standard, who can *correct* his action for the *reason* that it was erroneous, only of such a creature can one say that it follows and uses conventions. (Hacker 1987: 496)

It is for the very same reason that Hacker believes that even the talk of cerebral maps is nonsensical: Because maps are only maps of something if appropriate conventions exist. There simply is no such thing as representing a territory on a map without employing specific sets of conventions of representation including specific methods of projection (e.g. the Mercator projection).

So there are no representing maps without conventions of representation. There are no conventions of representation without a *use*, by intelligent, symbol-employing creatures, of the representation. And to *use* a representation correctly one must *know* the conventions of representation, understand them, be able to explain them, recognize mistakes and correct or acknowledge them when they are pointed out. Whether a certain array of lines is or is not a map is not an *intrinsic* feature of the lines, nor even a *relational* feature (that is, the *possibility* of a 1:1 mapping), but a *conventional* one (that is, the *actual* employment, by a person, of a convention of mapping). (Hacker 1987: 497f.)

Thus, one is forced to accept the conclusion that the idea of a language of the brain is literally nonsensical. There can be no meaningful symbols in the brain, because meaning presupposes the existence of conventions and conventions in turn imply the existence of a corresponding social practice. A „social practice“ of the kind required however, is *conceptually* impossible with respect to brain cells. The assumption that the brain employs a language or uses a system of symbols is therefore literally „inconceivable“.

2. At first sight this argumentation appears to be extremely plausible. And it indeed forms the core of a Wittgensteinian theory of meaning, which is shared by many. A closer look, however, will reveal that this argumentation is not quite as cogent. This is so because even the reference to a social practice cannot – at

least if one follows Kripke's reasoning concerning this point⁴ – provide grounds for the *normative* character of meaning. This is at least the way in which Paul Boghossian reads Kripke.⁵ Boghossian asks what the normative character of meaning consists in and answers:

Suppose the expression 'green' means *green*. It follows immediately that the expression 'green' applies *correctly* only to *these* things (the green ones) and not to *those* (the non-greens). The fact that the expression means something implies, that is, a whole set of *normative* truths about my behaviour with that expression: namely, that my use is correct in application to certain objects and not in application to others. ... meaningful expressions possess conditions of *correct* use. (Boghossian 1989: 513)

From this follows the sceptical problem for all theories of meaning.

Having a meaning is essentially a matter of possessing a correctness condition. And the sceptical challenge is to explain how anything could possess *that*. (Boghossian 1989: 515)

Correspondingly, Kripke's main argument against all theories which attempt to reduce meaning to natural properties of individual persons, and especially against the dispositional analysis of meaning, runs like this: None of the natural properties presented by these theories can account for the fact that expressions have conditions of correctness, and it is precisely because of this that all these theories as theories of meaning are doomed to failure.

At this point the Wittgensteinian brings into play rules which are grounded upon social practices and argue: Everything said so far is right, but what it shows is simply that meaning is not constituted through properties of *isolated* individual persons. The meaning of a linguistic expression only springs from the rules on which the use of the expression in question is founded. And these rules, in turn, result from a common social practice. But does this answer suffice? Can rules and can especially a social practice give better grounds for the conditions of correctness of a linguistic expression than the properties of individual persons?

Following Hart (1961: 54ff.) we can explain the fact that in a community there exists a rule *R* as follows:⁶

- (1) The members of the community rarely deviate from *R*.
- (2) If a member of the community deviates from *R*, then (s)he is exposed to sanctions from the other members of the community.

⁴ Kripke 1982.

⁵ Boghossian 1989.

⁶ See also von Savigny 1983: 34.

(3) These sanctions are – generally – accepted.

If this is so, then the fact that there exists a rule within a community consists only in the *dispositions* of the members of that community. And this in turn leads to the question: ‘In which way can the dispositions of a number of people provide better grounds for the conditions of correctness than the dispositions of an individual person?’

This is the reason why Kripke himself accepts the reference to the rules of a linguistic community only as a *sceptical* solution of the problem of meaning. A *substantial* solution is, he thinks, impossible. Nothing *in the world* can account for the normative character, i.e. the conditions of correctness, of linguistic expressions. Therefore, in a strict sense, the conclusion is inescapable that no linguistic expression has the *property* of having a certain meaning. Hence it is nonsensical to ask what this property consists in. The only thing we can do is to *describe* under which conditions we *ascribe* which meanings to which words, and perhaps ask why we do it this way rather than another.

Following this line we then find, according to Kripke, that in the ascription of meaning we actually *do* refer to actions and dispositions of members of linguistic communities. And, what is more, Kripke also holds – in common with many Wittgensteinians – that it simply does not make any sense, i.e. does not serve any intelligible purpose, to ascribe meaning to the utterances of an isolated individual person and that therefore our reference to social practices is not accidental but in a certain way inevitable.

However, if – in investigating the problem of meaning – the only thing we can do is to *describe* our practice of ascribing meanings to sentences and parts of sentences and to *explain* this practice, then there might be more alternatives available.

In this spirit I will explore in the following, whether there are not some good reasons after all, for the practice of many cognitive scientists who regard certain physical (e.g. neuronal) structures as representations with a certain meaning. If it should turn out that this in fact is so, this would in my opinion also show that speaking of a language of thought (or the brain) – notwithstanding the arguments of Hacker and others – has a perfectly intelligible sense after all.

3. However, I would like to begin with a concession. Hacker has made it very clear that according to our normal use of the word ‘language’ a language can only exist if there are beings who speak this language and that it can only be said of a being that it employs a language if it masters a certain broad range of

behavioral patterns.⁷ One of his arguments against the idea of a language of the brain was precisely that neither the brain nor parts of it *can* master such a behavioral repertoire. And in this he is certainly right.

A language of thought, therefore, can only exist if it is – in a certain way – radically different from all normal languages, for a language of thought, if it exists, is a language which is not spoken by anyone, nor understood by anyone – it is not even heard by anyone. (If some people talk as if the brain would speak or understand this language, then this mode of speech can only be meant metaphorically.) A language of thought is, as it were, a language which simply happens. Sentence tokens of this language just arise in the brain under certain conditions, are altered in accordance with certain formal rules and – together with more sentence tokens – cause certain actions. The sentence tokens need not be uttered in order to exist and have (causal) effects. All this happens, one is almost tempted to say, as if by itself. Given these conditions however, the question suggests itself: To what extent can one speak here of a language at all? This question is certainly justified and I am not absolutely certain whether it can be answered convincingly. However, I would like to begin tentatively with the following consideration: A language can first of all be simply conceived of as a system of structured sentences with combinatorial semantics. The sentences have a meaning (truth conditions) and this meaning depends in a systematic way on the meaning of its constituents. One can distinguish between sentence types and sentence tokens. Sentence tokens are physical structures concerning which one can tell which sentence type they realize. If one accepts this, one can perhaps agree with the following as well: If a number of physical structures exist in a system which can – with good reasons – be conceived of as tokens of certain sentence types and insofar also as having certain truth conditions, then there exists an internal language in the system. Perhaps, someone might claim that the term ‘language’ would be inappropriate in such a case and would instead prefer to speak of a system of internal representations. To this I would have no objections since systems of internal representations are all the cognitive scientist needs. And I am quite sure that no cognitive scientist ever took a language of thought to be something more than this. On the other hand,

⁷ In my opinion it is a very interesting question whether the whole behavioral range is *really* a necessary condition for the possession of a language, or whether we would not be inclined (or even forced) to attribute a language to beings who only possess a part of the skills Hacker mentions. Unfortunately I cannot pursue this question here any further.

Hacker's arguments, as he makes clear enough,⁸ are meant to count against systems of internal representations in the same way as against the idea that there could be more fullblooded languages in the brain. To opt for the former alternative, therefore, does not change the overall dialectical situation.

In the remaining sections I am going to argue for the thesis, that there really are good reasons for conceiving of certain systems in the way explained in the last paragraph (or that it is at least possible that there are some) and that therefore the idea of a language of thought (i.e., of systems of internal representations) in the sense described above is not at all nonsensical. The introduction to this will be a very general remark from the field of the philosophy of science.

4. If we try to explain and understand the of complex systems it is often not enough to take only the *physical stance*, as Dennett⁹ calls it. Often an adequate understanding is reached only when we also understand the *functional organisation* of these systems. That this is so becomes especially clear in the field of biology: There explanations are frequently given on the functional level alone while anatomical and physiological details are hardly mentioned. Let us take an example – for instance temperature regulation in the human body, which is explained in the textbook *Biological Psychology* by Birbaumer and Schmidt as follows:¹⁰

Thermoregulation can formally be viewed as a closed circuit regulatory system with a negative feedback loop. Body temperature is monitored by sensors, namely the thermoreceptors, which feed information into the central regulator. The latter checks whether the body temperature (the *actual value*) has deviated from its *desired value* and alters the control medium by sending *control signals* until feedback from the thermoreceptors signals that the mismatch has been compensated.

The body's core temperature is registered at different sites through temperature sensitive cells or sensory neurons, the surface temperature through thermoreceptors within and beneath the skin. The hypothalamus, especially the posterior hypothalamic area is likely to be the integration centre of temperature regulation. Central effector neurons control (probably via a chain of interneurons) the final control elements for the production and extraction of warmth (production of warmth, insulation of the body surface, sweat production and). They receive their afferent input from peripheral and central thermoreceptors. Cold receptors directly activate the effector neurons for the production of warmth and inhibit, via some interneurons, the final control elements for the

⁸ See for example Hacker's claim: „Nothing in the cortex constitutes a ‘symbolic representation’ of the creature’s environment.“ (Hacker 1984: 497)

⁹ The distinction between physical, functional and intentional stance goes back to Dennett 1971.

¹⁰ I'm abridging this description strongly.

extraction of warmth. Warmth receptors are wired in exactly the opposite way to the two types of effector neurons. (Birbaumer and Schmidt 1990: 117-121)

The almost exclusive use of functional vocabulary is obvious. Sensors, control media and feedback control systems are mentioned as often as integration centres, thermoreceptors and effector neurons. The only genuinely physiological concepts seem to be anatomical expressions like ‘posterior hypothalamic area’, and this is the case even though the story *could* be told completely also in purely physiological terms. But – apart from the fact that this story isn’t known to us in all its detail – *this* story alone wouldn’t satisfy us, because what we are really interested in is the question of how the body manages to maintain a relatively constant temperature under extremely differing conditions. And we only understand *this* if we realize that the physiological processes interact in the form of a feedback control system and therefore can be described with the help of the corresponding conceptual scheme. Functional concepts therefore, are brought in especially if one isn’t mainly interested in explaining individual physical states or activities, but in understanding how *successful* behaviour comes about, i.e. how a system manages to produce under the most varied conditions, behaviour which meets certain standards. We can sum this up as follows:

Thesis 1 *Often we can only explain and understand the successful of systems adequately if we proceed from the physical stance to the functional stance with regard to these systems.*

I’d like to add here a short remark concerning functional systems: In this context, it is important to note that properties such as being a sensor or being a final control element are not natural properties in the usual sense of the term. This means that we cannot ascribe concepts such as ‘sensor’ and ‘final control element’ – as opposed to concepts like ‘pyramidal neurons’ or ‘neuromuscular synapses’ – on the grounds of normal observable or measurable neurobiological characteristics. This is so because the applicability of these concepts to certain neuronal phenomena depends on whether these phenomena interact in such a way that a circuit pattern results which can be *interpreted* as a closed circuit feedback control system. To put it in a rather simple – though somewhat misleading – way: functional properties do not exist in the world, we *read* them *into* the world.

5. The example of thermoregulation, however, is a little too unspecific as to allow conclusions about the sense or nonsense of the idea of a language of thought. Another example might be a bit closer to the point – namely, the

example of a chess computer which Dennett has often used for the purposes of illustration.¹¹ For such an electronic device, it is possible – in principle at least – to explain every move in purely physical terms: One can ascertain how certain local states of silicon chips change through pressing certain letter or number keys; one can further deduce the sequence of the states these chips will go through after pressing ‘Enter’ from the circuit and the initial states of these chips. In the same way one can finally calculate which state will end this sequence and which of the diodes which make up the display will be lit. What can be achieved in this way, however, is only the explanation of certain concrete final states on the basis of knowing the concrete initial conditions. What cannot be achieved is an understanding of the mechanisms that enable the device to produce outputs which correspond to moves which are *plausible*, or even *successful*, in the relevant situation of the game.

Such an understanding can again only be reached if we move on from the physical to the functional stance. In this particular case, this amounts to analysing the *program* which underlies the behaviour of the chess computer. Because only then is it possible to conceive of what happens between input and output not just as a sequence of states of silicon chips. Only in the functional stance can we interpret certain local states of these chips as representations of possible configurations of pieces on a chess board. Only if we presuppose the functional stance can we describe the occurrences between input and output in a way which is almost familiar by now: The computer first calculates the representations of all possible successive configurations of the actual situation which would result from the moves possible for the computer; then it repeats this calculation for all the moves open to the opponent to respond to these, and again for the computer’s own possible moves to respond to the opponent, and so on until a certain number of moves and countermoves has been reached. The individual configurations are evaluated according to given criteria, before finally the computer produces, as an output, the move which leads to the configuration with the highest evaluation, taking into account optimal moves of its opponent.

This way of telling the story makes it possible, for the first time, to understand that our computer normally makes plausible, or even successful, moves, because it can be shown that the evaluative function underlying the choice of moves indeed results in plausible, or even good, moves under the conditions in question. If the computer in the end makes the move which leads

¹¹ First in Dennett 1971.

to the highest evaluation, its moves must, on average, be rather good ones. I say 'on average', because there are configurations which are objectively disadvantageous notwithstanding a high evaluation. If such a situation occurs, the computer often doesn't choose a particularly good move. This however, need not surprise us, because we know, of course, that the computer sometimes makes mistakes. So the description of what takes place between input and output, with the help of the program outlined above, is doubly helpful in explaining the behaviour of the computer: The description explains why the computer normally chooses good moves, and it also explains why it sometimes makes grave mistakes.

The example of the chess computer, as well as many examples of biological systems, show that we can often understand the of complex systems only if we proceed from the physical to the functional stance, and that this is particularly so if the in question is of a kind which, measured against certain standards, can be classified as successful. More important than this general point, however, is a point which comes to our attention if we take seriously the functional stance towards certain systems, e.g. towards a chess computer.

I have already mentioned that assuming the functional stance with respect to a chess computer amounts to analysing the program implemented by this computer. And this in turn means two things: firstly, we conceive of certain steps which take place between input and output as the execution of certain instructions, and, secondly, we reconstruct how the system organizes the sequence of these steps. The execution of a certain instruction normally consists in the production or manipulation of a certain data structure. This means that we can only conceive of certain physical processes as the execution of an instruction if we also at the same time view certain physical structures as data structures. With respect to the functional analysis of the chess computer, this means concretely: We can reconstruct the program which underlies its normally successful only if we conceive of certain physical structures within the system (the local states of certain silicon chips) as representations of possible configurations and of other physical structures of this kind as representations of evaluations. If we generalize this result we arrive at

Thesis 2 *The functional analysis of a system is in some cases only possible if one conceives of certain physical structures within the system as representations.*

Following Hacker's line of argument one might be tempted to object that the argumentation up to now does not take into account the fact that chess computers are artifacts which have, indeed, been programmed by their

manufacturers with a certain purpose. With respect to these artifacts, one can therefore say that they carry out programs and hence that within them there exists something like representations, because in this case there is someone – namely the programmer – who intends to represent certain configurations of chess pieces by means of certain physical structures. Representations without a person who uses them, however, would still be impossible.

This objection however, would miss the very point of my argumentation. Since this point is precisely that, with regard to some systems – *independent of their origin* – we *must* assume that there exist representations in them *if* we want to understand how the successful of these systems comes about. We therefore would have to describe chess computers in exactly the same way as I have explained above, even if they were to grow on trees.

And it can be easily shown that this explanatory strategy is, indeed, pursued in neurobiology. I remember vividly a discussion in the course of which I once asked the Göttingen physicist Manfred Schroeder which neuronal mechanisms are responsible for the localization of sources of sound. His answer began with the sentence: „Firstly the crosscorrelation of the signals of the two auditory nerves is calculated in the brain“. Another example of the same type can be found in J. Koenderink’s article „The Brain a Geometry Engine“; it is Koenderink’s central thesis that the best way to understand the mechanisms of the visual cortex is to take as a starting point the two dimensional intensity distribution of the light quanta which strike the retina and then to interpret the following neuronal processing as the calculation of the first, second and higher differentiation of this distribution.

... you may understand a large part of the structure of the front-end visual system as an embodiment of differential geometry of the visual field. ... Instead of the concrete ‘edge detectors’ and ‘bar detectors’, one speaks of the abstract first- and second-order directional derivatives. (Koenderink 1990: 125)

I cannot go into more detail here, but I hope it becomes clear even from these sketchy examples that in fact many neurobiologists take the functional stance in order to attempt to explain the amazing achievements of the brain, and that they go even further and try to reach explanations on the basis of the assumption that in the brain certain calculations really take place.

6. From this it is only a small step to my general conclusion: Just as it is necessary to assume that, within chess computers, there are representations of configurations and evaluations in order to understand how these devices succeed in producing successful moves, it may be necessary, with respect to other systems, to assume that there are sentence-like representations within

them, if one wants to understand what enables these systems to behave successfully. (In this context no more is meant by the expression ‘sentence-like representations’ than ‘structured representations with combinatorial semantics’.) This would – for example – apply to all AI systems, the problem solving of which is based upon automatic theorem proving, because we cannot adequately understand the of these systems without interpreting some of the processes taking place within them as *inference processes*. And inference processes are processes in which sentence-like representations are derived from sentence-like representations. That is to say, we cannot conceive of some processes within the system as inferential if we are not prepared to interpret some of the physical states within the system as sentence-like representations. Here – like in the example of the chess computer – it can be seen that the interpretation of processes has priority over the interpretation of states: Certain processes in a system cannot be adequately understood if we do not interpret certain states in a corresponding manner.

And now I think it is also clear under which conditions we are virtually forced to assume that there are sentence-like representations or symbols within certain systems, i.e. that these systems contain a language of thought. We are forced to assume this if we can only understand what underlies the successful of those systems if we interpret some of the physical processes within them as processes of production and manipulation of sentence-like representations. To sum this up in a last thesis:

Thesis 3 *The assumption of sentence-like representations is not only plausible, but in a certain sense unavoidable if we can explain the successful of a system only by means of the assumption that it is founded upon functional processes which can only be understood as processes of the production and manipulation of sentence-like representations.*

Speaking of sentence-like representations therefore is neither one of the little quirks of certain cognitive scientists, nor a habit which springs from a fundamental confusion. Rather it is a consequence which results from the attempt to understand the functional architecture of some systems underlying their successful .

By way of conclusion I want to emphasize strongly that Thesis 3 only formulates a *condition*. If this condition is satisfied, then we can say that within a system there exists a language of thought or a system of internal representations. This thesis however, does *not* imply that Fodor, or other cognitive scientists, are right in believing that intelligent can only be explained within the symbol processing paradigm. However, my purpose was not to

defend this paradigm, but rather to save it from the charge of conceptual confusion.

References

- Birbaumer, N. and Schmidt, R.F. 1990 *Biologische Psychologie*, Berlin: Springer.
- Boghossian, P. 1989 „The Rule-Following Considerations“, *Mind*, 98 507-549.
- Dennett, D. 1971 „Intentional Systems“, *Journal of Philosophy* 68, 87-106.
Reprinted in D. Dennett, *Brainstorms*. Montgomerly, VT.: Harvester, 1978, 3-22.
- Fodor, J.A. 1975 *The Language of Thought*, New York: Thomas Y. Crowell.
- Fodor, J.A. 1978 „Propositional Attitudes“, *The Monist* 64, 501-523. Reprinted in J.A. Fodor, *Representations*, Cambridge, MA: MIT Press 1981, 177-203.
- Fodor, J.A. 1981 „Introduction – Something on the State of the Art“, in J.A. Fodor, *Representations*, Cambridge, MA: MIT Press, 1981, 1-31.
- Fodor, J.A. 1987 *Psychosemantics*, Cambridge, MA: MIT Press.
- Hacker, P. 1987 „Languages, Minds and Brains“ in C. Blakemore and S. Greenfield (eds.), *Mindwaves: Thoughts on Intelligence, Identity and Consciousness*, Oxford: Blackwell, 1987, 485-505.
- Harman, G. 1973 *Thought*, Princeton: Princeton University Press.
- Hart, H.L.A. 1961 *The Concept of Law*, Oxford: Oxford University Press.
- Hyman, J. (ed.), 1991 *Investigating Psychology*, London: Routledge.
- Koenderink, J.J. 1990 „The Brain a Geometry Engine“, *Psychological Research*, 52, 122-127.
- Kripke, S. 1982 *Wittgenstein on Rules and Privat Language*, Oxford: Blackwell.
- von Savigny, E. 1983 *Zum Begriff der Sprache*, Stuttgart: Reclam.
- Stillings, N.A. et al. 1987 *Cognitive Science: An Introduction*, Cambridge, MA: MIT Press.