Rationality as an explanatory principle in linguistics

ESA ITKONEN
(Helsinki)

1. Concerning the general definition of rationality

As general principles rationality and correctness, which may be subsumed under the superordinate concept of normativity, must be contrasted to factuality understood as spatio-temporal existence. The natural sciences investigate only that which is factual whereas some human sciences investigate either that which is rational or that which is correct. Claims about what is factual, rational, or correct share the property of not being falsified by what is, respectively, nonfactual, irrational, or incorrect. The difference between normativity and factuality consists in the fact that what is (ir)rational/(in)correct may or may not occur in space and time, whereas what is factual occurs in space and time, and what is nonfactual does not. Hence a theory of natural science or of empirical sociology is falsified if what it predicts to be factual is nonfactual, as shown by something that *is* factual, i.e. some *un*predicted spatio-temporal occurrence [1]. A theory of philosophy, of logic, or of autonomous linguistics is falsified if what it predicts to be rational/correct is irrational/incorrect; but here the reference to spatio-temporal occurrences is doubly irrelevant. Suppose that a sentence s_1 predicted by a grammar to be correct is, in

[1] Falsification of theories should be confused neither with their apparent falsification nor with their rejection. First, the history of science shows that even when a theory has been contradicted by evidence, it is rational to suspend the judgement for some time, since the evidence may turn out to be faulty in one respect or another. Second, even when a theory is genuinely contradicted (or falsified) by evidence, it is rational not to reject it until a better theory has been invented. These remarks suffice to show that Lakatos' (1970) and, following him, Botha's (1978) strictures against 'falsificationism' are misconceived. Nevertheless, I do not accept Popper's exclusive emphasis on falsification. Proving that a theory A is better than a theory B is more fundamental than proving that B is false and A is not (i.e not yet), because doing the latter is only one way, even if the most important way, of doing the former.

fact, incorrect while a structurally similar sentence s_2 predicted to be incorrect is, in fact, correct. If this is the case, the grammar has been correspondingly falsified, and yet it is possible that s_1 has occurred in space and time whereas s_2 has not, which shows that reference to space and time is indeed doubly irrelevant in grammatical falsification (for details, cf. ITKONEN 1978: 9.0-11.0).

In the end of the preceding paragraph it became evident that we can know what is rational/correct without simultaneously observing what occurs in space and time or remembering what has done so. (This act of knowledge pertaining to normativity is customarily called 'intuition'.) In other words, we are able to construct mental representations of rational actions or correct sentences and to recognize them as rational or correct, even if such actions or sentences have (as far as we know) never been exemplified in space and time. It must be emphasized, however, that this ability has not developed in a vacuum; we obviously learn the concepts of rationality and correctness as part of the process of becoming members of our community, and this process continuously involves observing both actions and *re*actions to them. But somewhere there seems to occur a 'leap' as a result of which the rules of rational/correct behavior are grasped, as is shown by the well-known philosophical truth that knowledge of what ought to be done cannot be reduced to (observation of) what is done.

As a consequence, I do not think that the reductionist attitude vis-à-vis normativity is justified. Today there are two common types of attempt at reduction. First, there is the genetic approach, or the attempt initiated by GRICE to reduce (part of) normativity to non-normativity by showing how the conventional communicatory act might have evolved from the general non-conventional one (cf. e.g. BENNETT 1976). One can agree that Gricean accounts are plausible cases of stylized history and yet regard them as rather pointless. The qualitative leap from non-conventional to conventional still has to occur somewhere which means that it has not been explained away. (Notice that similar leaps must have occurred between the formation of the first living cell and the emergence of non-conventional communication.) Second, there is the psychophysiological approach, or the attempt to reduce normative to non-normative by speculating how rational/correct behavior might be deduced from general psychological and ultimately physiological regularities. I have criticized this approach at some length in ITKONEN (1978: 7.0). Here I only wish to point out that scientific argumentation or the wish to convince one's interlocutors necessarily presupposes the concept of rationality, even when one is trying to explain rationality away, which makes such a programme contradictory (cf. HOLLIS 1977, chap. 7). A physiological explanation of human rational behavior can be given only by more-than-human beings in the same way as we can conceivably give a physiological explanation of the psychology of frogs, for instance.

The distinction between rationality and correctness may be quite simply illustrated by the two different types of answers given to the question «Why did you say 'The window is broken'?» One answer would refer to the purpose of uttering the sentence, e.g. «I said 'The window is broken' in order that someone would fix it», whereas the other would refer to rules of English, e.g. «I said 'The window is broken' because the past participle of *break* is *broken*, and not *breaked*». In other words, rationality is a matter of choosing means adequate to ends, while correctness is a matter of following well-established rules. It must be emphasized that contrary to the prevailing opinion, rules of language, which must not be confused with rules of grammar, *can* be brought to the level of consciousness (cf. COSERIU 1958/1974: 49-51; for discussion cf. ITKONEN 1978: 5.3).

2. *The different roles of rationality*

At the level of research, rationality plays an identical role in all sciences[2]: the goal of the scientist is to bring order into chaos, and he uses means which he considers adequate to attaining his goal. Even if the goal is the same, the means of course differ widely, depending on the science in question. Thus it is at the level of research objects, or of theories invented to account for research objects, that rationality plays different roles in different sciences. It is self-evident that at the level of research, object rationality has no role to play in the natural sciences. In the human sciences it is possible to distinguish between (at least) three ways in which the concept of rationality becomes part of the theory. At the same time we will have reason to distinguish between three different but closely related meanings of 'rationality'.

First, we may start from a set of experimental or observational data of human behavior and try to apply a straightforward natural-science model to explain them. However, it may turn out that this method fails to produce illuminating results, i.e. fails to bring order into chaos. Now what is felt to be illuminating or unilluminating, is largely a matter of personal taste, and therefore it is possible that some people who follow the natural-science approach are entirely satisfied with their results. But even if one is not satisfied, one still has two options. Either one tries to intensify the imitation of the natural sciences or one gives up the natural-science model and adopts instead the (teleological) model of rational agent. The latter option has been adopted or is being adopted in such disparate sciences as psychotherapy (cf. ITKONEN 1978: 2.3), psychology of perception (cf. NEISSER 1967), and diachronic linguistics (cf. Sect. 3 below). It is to be noted that what is needed here is a concept of *unconscious* rationality. It is also clear that at least in psychotherapy we must extend

[2] I use the term 'science' in the sense of 'Wissenschaft'. Hence both physics and logic are sciences; even philosophy of science is itself a science.

the everyday concept of rationality to cover cases which might normally be considered irrational: in an unbearable 'no-win' situation, neurosis may appear as the only 'rational' solution. — In all cases discussed in this paragraph it is possible to refuse to adopt the model of rational agent (although I personally think that it should definitely be adopted).

Second, in sociological analysis we start from correlations between observable variables and postulate theoretical variables to explain the correlations[3]. This method is analogous to the standard natural-science method, except for one thing: It is required *a priori* that correlations between variables be *understandable*; it is in fact the point of postulating theoretical variables to establish the understandability where it is lacking. That is, we do not, of course, know *a priori* which correlations we will find, but we do know *a priori* that we will not accept them unless they are or can be made understandable. Consider LABOV's (1972) analysis of vowel centralization on Martha's Vineyard. This linguistic variable was found to correlate with the social variable 'middle-aged fisherman', but the correlation could not be accepted as such, because it means nothing, i.e. it seems just a coincidence. The correlation was accepted only when it was explained or made understandable by the intervening psychological variable 'positive attitude towards staying on the island'; and this variable was, of course, postulated (and later operationalized) precisely because it made the correlation understandable or revealed its *meaning*. (This psychological variable is, of course, a causal one; but causes of actions are intentions and motives, often unconscious, and they are understood roughly in the same way as meanings of sentences.) Precisely the same remarks apply to any standard sociological analyses, starting from DURKHEIM's paradigm-creating study of suicide and his theoretical-explanatory concept of 'anomie'. By contrast, no similar *a priori* requirement of understandability can be imposed upon the theoretical concepts of natural sciences. For instance, there is surely nothing understandable or meaningful in the concepts of 'wave-particle' or 'curved space'. — Equating 'rational' with 'understandable' again amounts to an extension of the everyday concept of 'rationality', an extension similar to, even if less drastic than, the one carried out in psychotherapy. I do not think that the existence of this second type of rationality can be sensibly denied.

Third, (our knowledge of) the concept of rationality can be taken directly as the object of analysis. This is true of the game theory which, as a formalization of decisions made under certainty, uncertainty, or risk, constitutes a nonempirical theory of rational behavior[4]. The theory of speech acts is another example of the same type of theory; that is, it is

---

[3] To be sure, sociologists, including sociolinguists, much too often remain at the first, non-theoretical stage of this research strategy.

[4] It is possible, however, to utilize experimental evidence to constrain the primarily nonempirical approach of the game theory in the direction of increasing empiricalness; cf. Diesing (1972, chap. 4).

a philosophical theory based on intuitive knowledge about how to act, i.e. speak, rationally within the normative space defined by the rules of language. Today's psycholinguistic theories of speech production have incorporated the speech act theory (cf. Sect. 4). — The third role of rationality is comparable to the role of correctness in autonomous linguistics or the role of validity in formal logic. It would be absolutely impossible to deny the existence of this type of rationality because here rationality, instead of being a hypothesis arrived at on the basis of spatio-temporal data or of constituting a precondition upon (correlations between) such data, simply *is* the data. This type of rationality, being a potential object of conceptual analysis, is necessarily capable of being brought onto the level of consciousness.

3. *Diachronic linguistics*

In what follows, I shall concentrate on morphological change. In modern diachronic linguistics the natural-science approach has been adopted nearly universally[5]. That is to say, linguists are looking for universal regularities of change which could be described by sentences of the type '(x) (Fx ⊃ Gx)'; together with suitable antecedent conditions *Fa*, a regularity of this kind would explain the change *Ga*. In the field of diachronic morphology at least, this search for regularities has so far been futile.

I wish to argue that this state of affairs does not result from linguists' temporary incapacity to hit upon the right regularities but rathed reflects the true state of affairs. Both morphological change and morphological resistance to sound change[6] are of *teleological* character. The ultimate goal, which may be exemplified by an unlimited variety of lower-level goals, is to maintain the principle 'one meaning — one form' (cf. PAUL 1920/1975:227) or to restore it once it has been violated by sound changes. This goal can be used afterwards to explain morphological changes, but it cannot be used to predict them, which means that there *are no* universal regularities of morphological change. In this respect teleological morphological change resembles all rational behavior. In fact the teleological explanations here at issue are in my opinion a type of *rational explanation*[7]. More precisely, what we have here is *unconscious* rationality, whose existence has been independently assumed in several fields of study (cf. Sect. 2). The basic form of rational explanation is so-called practical inference (cf. VON WRIGHT 1971) which as a model of explanation is characterized by the fact that it contains no reference to regularities. — The view of diachronic-linguistic explanation as a type of

[5] Coseriu (1958/1974) and Anttila (1972) are here important exceptions.

[6] In what follows, 'morphological change' will stand for both of these two cases.

[7] There is also a type of teleological explanation which makes reference to (teleological) regularities, but this is not what I have in mind here.

rational explanation was put forth by COSERIU (1958/1974, esp. p. 57 and 158-59), and is also discussed in ITKONEN (forthcoming).

It is worth emphasizing that rational behavior is never *entirely* unpredictable. In each situation rationality defines a set of possible courses of action and excludes a much larger set of impossible, i.e. irrational, courses of action [8]. Determining the (probable) limits of rationality is an important task which has not yet been carried out in diachronic linguistics.

What I have been saying so far might be countered by pointing out that the requisite regularities of morphological change need not be universal but merely statistical, and that the appropriate model of explanation is accordingly statistical, and not deterministic, in character. This objection is insofar justified as it is certainly possible to gather such statistical data concerning morphological change which condenses into one regular pattern or another. Such a programme contains one drawback, however: The explanatory force of statistical explanations is weaker than that of deterministic explanations; the latter may at least tentatively explain why something happened, whereas the former can at most explain why it was to be expected that something would happen. In other words, in statistical explanations there is always an element of chance which weakens the tie between that which explains and that which is explained. This is, of course, no argument against using statistical explanations if statisticalness is not just a methodological artifact, or a result of the researchers' ignorance, but belongs to the ontology of the research object, as seems to be the case e.g. in subatomic physics. But there is reason to believe that rational behavior does not belong to the same category as physical behavior, whether subatomic or not. By concentrating on the specific context in which an action, whether individual or collective, has come about, it is possible to reveal its coming-about as more necessary and less dependent on chance than could be done by means of explanations resting on statistical regularities (cf. HARRÉ & SECORD 1972, chap. 7, esp. p. 133).

From the previous paragraph it is evident that the type of explanation I am recommending for diachronic linguistics (above the phonological level) comes rather close to the traditional notion of historical explanation, according to which, historical phenomena are to be explained by reference to their socio-cultural context and to the personalities of the actors, and not by reference to some eternally and ubiquitously valid regularities. This position has been called into question by the wellknown 'POPPER-HEMPEL thesis', which claims that historical explanations cannot avoid ultimately making use of precisely such regularities. However, the POPPER-HEMPEL thesis rests on a blind and unquestioning faith in the natural-science model and is entirely unsupported by historians' actual

[8] It must be added that there does not seem to exist just one single concept of rationality. The game theory at least recognizes several concepts of rationality, none of which seems intuitively superior to the others.

descriptive practice. DONAGAN (1966) points out that *not a single* historical explanation satisfying the standards of the POPPER-HEMPEL thesis has *ever* been proposed. Such sentences as purport to refer to universal regularities of social behavior are either maxims of rational behavior and hence analytically true or else quite obviously false. Neither type of sentence can serve as a basis for empirical explanation. This being the case, the POPPER-HEMPEL thesis may safely be dismissed.

4. Psycholinguistics

Psycholinguistics is primarily concerned with uncovering the mechanisms underlying speaking and understanding. Speaking is a form of rational behavior, and therefore the theory of speaking must be just an application of the general theory of action (cf. ITKONEN 1972), with the proviso that the conventionality of language cannot, at least in my opinion, be explained away (cf. Sect. 1). It is convenient to divide acting, including speaking, into two components, viz. planning and execution (cf. CLARK & CLARK 1977, chap. 6). The component of planning constitutes the realm of rationality; execution may or may not proceed as planned. Understanding is normally an automatic process and therefore cannot be regarded as rational; its relation to rationality consists in the fact that it ideally means recovery of the plan 'behind' the action. Modern cognitive psychology considers planning and understanding as internal, mental processes only contingently related to the external reality. This view cannot be accepted, for two reasons. First, the goals for action are provided exclusively by the agent's socio-cultural environment. Second, understanding is devoid of content, if it cannot be distinguished from misunderstanding; but this distinction can be made only on the basis of social public criteria. The same holds true of planning. These remarks are in an obvious way related to WITTGENSTEIN's refutation of private languages (cf. ITKONEN 1978: 4.0).

It is generally agreed today that psycholinguistic theories cannot be constructed without recourse to so-called external evidence, i.e. evidence not based on intuition alone. It is in fact the pervasive mistake of transformational grammar that it confuses the concepts of autonomous linguistics and psycholinguistics by automatically ascribing psychological reality to formalizations of intuitive knowledge. These remarks apply to the psychological theories of linguistic levels from phonology to semantics. However, if one wishes to construct *large-scale* theories of speech production and understanding, no systematic external evidence is available, and therefore one cannot help relying on intuitive knowledge about acting in general, and speaking in particular[9]. GRICE's and SEARLE's theory

[9] The reason for this state of affairs will be discussed below.

of speech acts is based on precisely that kind of knowledge, and therefore it is only logical that CLARK & CLARK (1977) should have incorporated it as part of their 'planning component' (cf. above). This component is further divided into three subcomponents, within each of which the speaker has to make decisions about what it is rational to say, given what it is correct to say.

The upshot of the previous paragraph may seem somewhat surprising. Psycholinguistic theories are first and foremost *causal* theories: they mean to uncover those mechanisms which bring about speaking and understanding. It seems undeniable that psychological causes for the most part lie under the level of consciousness. And yet I claimed in the previous paragraph that (large-scale) psycholinguistic theories can be based on intuitive knowledge, which clearly seems to imply that psychological causes are, after all, open to conscious inspection. Moreover, I criticized transformational grammar for not distinguishing between autonomous linguistics and psycholinguistics, but now I appear to commit the same mistake: In Sect. 1 I defined autonomous linguistics as that type of linguistics which relies exclusively on intuition[10] but —to repeat— I stated in the previous paragraph that psycholinguistic theories can, and even must, be based on intuition.

This puzzle can be solved by distinguishing between two types of causal relevance. We must be clear as to which of the two following questions we wish to answer: «Which conceptual distinctions must people have internalized?» and «How have they internalized them?» Distinguishing in this way between the *what*-question and the *how*-question calls for two comments. First, the former question is the logically primary one; it is only when it has been answered that one can try to answer the latter. Second, the former question expresses conceptual necessity, whereas the latter speaks about contingent, causal relationships. The difference between the questions is parallelled by the difference between the respective methods by which they are answered. The former question is answered by (nonempirical) analysis of intuitive knowledge, whereas the latter is answered by (empirical) investigation of external evidence.

KANT's categories may be mentioned as examples of distinctions which causally influence people's thinking and acting in spite of the fact that they have been arrived at as a result of purely philosophical analysis. It is clear, for instance, that the category of 'quantity' and its three subcategories 'unity', 'plurality', and 'totality' are part of our everyday thinking, which means that we must have internalized this trichotomy. In just the same way a grammarian can with absolute precision state e.g. the distinction between active and passive sentences which English speakers must have internalized. But in both cases the question how the internalization is actually implemented remains entirely open. For instance, the grammarian

[10] In a more thorough analysis this claim has to be modified somewhat; cf. Itkonen (1977).

is not competent to decide whether the distinction between active and passive has been internalized by making one derivable from the other or by memorizing each separately.

Both dynamic causes (or 'triggering conditions') and static causes (or 'standing conditions') contribute to bringing about an event. If we apply this distinction to the study of speaking, we notice that internalizations of rules of language constitute the permanent system which is from time to time put into motion by rationality. We further notice that intuition about correctness, i.e. rules of language, is able to identify (part of) the static causes and hence to answer the *what*-question, but is not able to uncover their actual manner of functioning or to answer the *how*-question. Intuition about rationality, in turn, is able to identify (part of) the dynamic causes. It is interesting to note that when we are discussing dynamic causes open to conscious inspection, or conscious motives of actions, the distinction between the *what*-question and the *how*-question seems to disappear. When we are consciously striving after a goal and consciously choose a certain (sub)action as a (sub)means leading to the goal, we are not only identifying the (conscious) causes of our action [11], but we are also identifying the actual manner in which the causes are functioning. This is why no external evidence is needed in constructing large-scale theories of speaking (or understanding). The only qualification that has to be made concerns the possibility that the rationality of some actions is only apparent, and that they can be genuinely explained only by recourse to unconscious causes [12]. It is true, of course, that even genuinely rational actions must be somehow rooted in man's psycho-physiological make-up.

The model of rational agent is incompatible e.g. with CHOMSKY's (1976) conception of linguistics: On the one hand, he considers linguistics as part of cognitive psychology; on the other, he denies any difference between linguistics and physics; but cognitive psychologists imitating the example of physics cannot make use of the model of rational agent. By contrast, FODOR (1976) admits the necessity of the rationality assumption, but he does not seem to realize that he, by so doing, is breaking with the entire methodological tradition of transformational grammar (cf. ITKONEN 1978: 3.6).

5. *Answering the standard objection against rational explanations*

Even if we are able to determine what it is rational for a person A to do in a given situation (and normally we are not able to do so), it does

[11] I use 'cause' as a superordinate term covering both traditional 'causes' and traditional 'reasons'. By the use of this terminology I wish to show what is common to traditional 'causes' and 'reasons' without, however, wishing to obliterate the distinction between them.

[12] Assuming that the rationality of *all* actions is only apparent leads to a logical contradiction; cf. Sect. 1.

not follow that A will do it, because people may behave irrationally. Therefore, HEMPEL (1965: 10.3) and STEGMÜLLER (1974: 6.7) claim, rational explanations are not genuinely explanatory.

HEMPEL and STEGMÜLLER are in fact only repeating the truth that 'ought' does not entail 'is' (cf. Sect. 1). This conceptual point, though valid, does not discredit the use of rational explanations. In Sect. 4, I stated that rationality is a causally effective force. Therefore when we explain an action by revealing its rationality, we are giving a genuine, causal explanation. The only difficulty here is that, as a cause, rationality does not operate uniformly as physical causes do, and therefore we are not able to determine with certainty when the agent is rational and when not. This fact, namely that the agent may or may not be rational, adds to the uncertainty created by the fact that even when he *is* rational, he is as a rule free to choose from among several equally possible courses of action.

As a consequence, in addition to saying that the action to be explained was the rational thing (or, more realistically, one of the rational things) to do, rational explanations also contain a premise stating that the agent was in fact rational. HEMPEL and STEGMÜLLER claim that adding this premise transforms rational explanations into standard empirical explanations, because the explanandum-sentence may now be deduced from the premises, one of which is a general statement about what all rational agents do. But this is a *non sequitur* because far from referring to an empirical regularity, such a statement expresses an analytically true maxim of rational behavior (cf. the end of Sect. 3). Analytic truths about rationality cannot be strengthened by noting that they hold of *every* rational agent. In a similar vein, the analytic truth of the particular statement «If Bill is a bachelor, he is unmarried» cannot be strengthened by the analytic truth of the universal statement «For all x, if x is a bachelor, x is unmarried».

The only thing left is for HEMPEL and STEGMÜLLER to ask what kind of tie connects the explanans and the explanandum of rational explanations, if it is not logical deducibility as in standard empirical explanations. The answer is simple enough: it is the tie which connects the explanans and the explanandum of practical inferences (cf. VON WRIGHT 1971). When the action to be explained was *the* rational thing to do, or a necessary condition for attaining the goal, the connection is felt to be very tight or explanatory. When the action was one of the rational things to do, or (part of) a sufficient condition for attaining the goal, the connection is less tight. In many cases, however, this is the closest we can get to explaining actions.

BIBLIOGRAPHY

ANTTILA, R. (1972), *An introduction to historical and comparative linguistics.* New York.
BENNETT, J. (1976), *Linguistic behavior.* London.
BOTHA, R. (1978), «Protecting general-linguistic hypotheses from refutation». *Stellenbosch Papers in Linguistics* 1.
CHOMSKY, N. (1976), «Problems and mysteries in the study of human language». In: KASHER, A. (ed.), *Language in focus.* Dordrecht-Holland.
CLARK, H. and E. CLARK (1977), *Psychology and language.* New York.
COSERIU, E. (1958/1974), *Synchronie, Diachronie und Geschichte.* München.
DIESING, P. (1972), *Patterns of discovery in the social sciences.* London.
DONAGAN, A. (1966), «The Popper-Hempel theory reconsidered». In: DRAY, W. (ed.), *Philosophical analysis and history.* New York.
FODOR, J. (1976), *The language of thought.* Hassocks.
HARRE, R. and P. SECORD (1972), *The explanation of social behavior.* Oxford.
HEMPEL, C. (1965), *Aspects of scientific explanation.* New York.
HOLLIS, M. (1977), *Models of man.* London.
ITKONEN, E. (1972), «On Grice's, Strawson's and Searle's concept of meaning». *Ajatus* 34.
— (1977), «The relation between grammar and sociolinguistics». *Forum Linguisticum* I: 3.
— (1978), *Grammatical theory and metascience.* Amsterdam.
— (forthcoming), «Short-term and long-term teleology in linguistic change». In: MAHER, P. (ed.), *Papers from the IIIrd international conference on historical linguistics.* Amsterdam.
LABOV, W. (1972), *Sociolinguistic patterns.* Philadelphia.
LAKATOS, I. (1970), «Falsification and the methodology of scientific research programmes». In: LAKATOS, I. and A. MUSGRAVE (eds.), *Criticism and the growth of knowledge.* London.
NEISSER, U. (1967), *Cognitive psychology.* New York.
PAUL, H. (1920/1975), *Prinzipien der Sprachgeschichte.* Tübingen.
STEGMUELLER, W. (1974), *Probleme und Resultate der Wissenschaftstheorie und Analytischen Philosophie, vol. 1: Wissenschaftliche Erklärung und Begründung.* Berlin.
VON WRIGHT, G. (1971), *Explanation and understanding.* London.