# THE CENTRAL ROLE OF NORMATIVITY IN LANGUAGE AND IN LINGUISTICS*

Esa Itkonen

**Abstract**
'Any natural language consists of rules which are inherently social and normative.' It is the purpose of this chapter, first, to establish the truth of this claim; second, to show that it is significant or non-trivial; and third, to explore its many ramifications.
Keywords: normativity, intersubjectivity, sociality, psychologism

## 1. **The social nature of language and meaning**

The word 'language' can of course be used in many different senses, but it is reasonable to assume that one sense may be primary. Thus, when we speak of e.g. 'English', what kind of entity is it that we mean by this word? More specifically, is this entity social or non-social (in the sense of individual-psychological)? The 'common-sense' answer is that it is a social entity. It goes without saying (or so it seems) that e.g. a dictionary of English is about something that it *common to* or *shared by* all speakers of English, and whatever has these characteristics must be social by definition. But the 'scientific' answer (e.g. Chomsky 1965) is generally taken to be that linguistics is part of cognitive psychology, which entails that e.g. English is, at least primarily, an individual-psychological (and not a social) entity. I shall argue in this chapter that, on this particular issue, common-sense is right and science is wrong.

### *1.1 The private-language argument*

The primarily social nature of language can be shown in different ways. I have always preferred to rely on Wittgenstein's so-called private-language argument, or PLA for short. PLA has spawned a huge number of publications, among which Saunders & Henze (1967) still stands out. Considered with all its ramifications, PLA is anything but simple. A 'minimalist' version of it will be presented in what follows (but see also Itkonen 1978: 91–113, 2003b: 120–125).

PLA directs itself against the dominant tradition of Western philosophy, a tradition equally represented by Descartes, Hume, and Kant. According to this ('Cartesian') tradition, public things and qualities are based upon or reducible to subjective experiences, which constitute the 'rock bottom' of knowledge. Moreover, knowledge of *other minds* is supposed to be gained on the basis of the 'argument from analogy': When I perceive that bodies (constructed out of my sense-impressions and) resembling mine behave under similar circumstances in

the same way as my body does, I may infer with a high degree of probability that these bodies are possessed by minds which think and feel in ways similar to mine.

To start with, the incoherence of the Cartesian position may be demonstrated by a simple conceptual argument. The Cartesian ego, expressed as *I* or *me,* is supposed to be prior to other persons. But, just as there can be no 'left' without 'right', there can be no 'I' without 'you' and 'we': "If as a matter of logic you exclude other people's having something, it loses its sense to say that you have it" (Wittgenstein 1958: §398).

More elaborately, the Cartesian position may be reformulated in linguistic terms, as follows. Since knowledge of the intersubjective or public world is supposed to be based on subjective or private experiences, the ordinary intersubjective or public language must — or could — have been preceded by a subjective or private language. Such a language is *private* in the twofold sense that it refers to subjective experiences and its rules are known to one person only.

This explains why the refutation of Cartesianism is generally called PLA. Wittgenstein (1958: §§ 243–277 and passim) argues that if a person constructs a private language and *consciously* tries to follow its (private) rules, he cannot know whether or not he has made a mistake. Because the notions of language and rule presuppose the possibility of making a mistake, there can be no private language: "The test of whether a man's actions are the application of a rule is ... whether it makes sense to distinguish between a right and a wrong way of doing things in connection with what he does" (Winch 1958: 58).

Presented in outline, PLA goes as follows. Suppose that I am at this very moment going to (consciously) use some word X of my own private language. My use of X, i.e. what I mean (or intend to mean) by X, is based on my particular memory of how I have decided to use X, or how I have used X in the past. Maybe I wish to check this memory to make sure that I am not mistaken. But the only check I can rely on is the *same* memory; and of course it is no independent or *genuine* check — in fact it is *no* check (or basis for testing) at all. Therefore any 'private' rule-application that seems correct to me will be correct, which means that the notion of a private rule-application, and thus of a private language, 'dissolves' (cf. the Winch-quotation above). Written documents, for instance, do not get me out of this circle, because now the question arises whether I remember correctly the meanings of the written 'private' words. (Notice that on this reading of PLA the exact nature of the referent — thing or sense-impression? — is no longer of decisive importance.) Kenny (1973: 192–193) presents this argument exceptionally well.

Genuine checks are provided only by *other* people's memories, and more generally by their intuitions about the correct use of (public) language. Of course, there is no guarantee that these are always trustworthy. But at least they offer the *possibility* of genuine testing; and possible testing is certainly preferable to

impossible testing (represented by exclusive recourse to my own memories or intuitions). This is nothing but the requirement of *objectivity* (in the sense of 'intersubjectivity'), which is the cornerstone of scientific thinking.

Some readers may still remain unconvinced. Therefore, to further clarify the issue, let us deal with a concrete counter-argument which has kept reappearing in essentially the same form from the mid-50s onwards. Suppose that I formulate a 'private' rule according to which what is now called *blue* ought to be called *mlue* by me. I paint a blue patch on a piece of paper and write *mlue* under it, and on future occasions I will use this device to make sure that I am indeed following the rule correctly. Have I not proved that the notion of a private rule is a viable one?

The answer is 'No', and here are some of the reasons why. Taken together, the blue patch and the word *mlue* constitute a 'picture'. When composing this picture, I may have thought that its meaning is self-evident, i.e. that it can be interpreted in one way only. But this is wrong. One of Wittgenstein's basic insights is that every picture or image can be interpreted in an infinite number of ways (and this is also true of *mental* images; cf. Blackburn 1984: 45–50, Heil 1992: 25–30). On the next occasion when I look at the picture, I may mistakenly think that the rule was meant to be *not* to say *mlue* when seeing something blue; or I may think that the blue patch was meant to remind be that I should check whether in any of the world's languages 'blue' is called *mlue*; and so on.

In other words, the human memory is notoriously fallible. It would be preposterous to assume that I am the only person in the world whose memory happens to be absolutely infallible. Now, what is true of memory, is true of intellectual capacities more generally. Human beings may succumb to any kinds of aphasia, delusion or insanity. Today, with the well-documented spread of Alzheimer's disease, this has become a near-certainty: everyone of us, unless released by a timely death, will become (more or less) insane. Let us keep this in mind, when we now return to the explication of PLA.

Realizing the ever-present possibility of multiple interpretations, I may now wish to secure the unambiguous meaning of the picture by adding an explicit written instruction. If I use my own private language, the instruction will look something like this: "zmosh # glaark * mlue". But nothing can guarantee that I will remember the meanings of these private words correctly, and if I attempt to avert this danger by further amplifying the instruction, infinite regress will ensue. If, on the other hand, I use English, the instruction will look like this: "I ought to say *mlue* whenever I see this colour!" But now I am cheating because my supposedly 'private' language is based on a public one. More importantly, however, this does not help me at all, because now any of the forms of human frailty alluded to above may attack me, either one by one or jointly. Perhaps I am colour-blind, but just do not know it; or perhaps I have become insane and think that, when looking at a blue patch, I am looking at my face in a mirror; or perhaps the moment when I lose the mastery of English has already arrived (but I just do not know it) and I either

fail to understand the instruction or think that it says that I should go wash my teeth; and so on. The upshot is that my rule-following behavior needs checking by others. This is not fool-proof either. (Perhaps everyone is insane.) But at least it provides the possibility of genuine checking, which my private memory and understanding can*not* provide.

Barresi & Moore's (this volume) Intentional Relations Theory can be thought of as an empirical equivalent of PLA. In their requirement that, for a psychological concept to come about, "both the first-person, 'inner' aspect and the third-person, 'outer' aspect" are equally needed, they reproduce the insight that "[A]n 'inner process' stands in need of outward criteria" (Wittgenstein 1958: §580).

Wittgenstein assumes that in language, like in any other social institution, we are, or may become, *conscious* of the rules we either follow or break. Attempts to dispose of PLA are often based on redefining 'private language' as '*unconscious* psychological structure', which makes it self-evidently true that everybody has his or her 'private language'. But the redefinition is unjustified, in the first place. Just as well one might decide to call the internal structure of individual atoms their 'private languages'. — The interdependence of normativity and consciousness has been explored in an illuminating way by Zlatev (2006).

A useful up-to-date explication of what it is to follow a rule (of language) is provided by the doctrine of 'response-dependency' or 'response-authorization' (cf. Pettit 1996: 195–204, Itkonen 1997: 58–60, 2003b: 126–130, 165–168, Haukioja 2000).

Nothing of what precedes entails in any way that language is *exclusively* social in character. Language has of course both a psychological and a biological aspect or, if you like, 'substratum'. What the preceding discussion is meant to establish is that language is *primarily* a social entity.

### *1.2 Semantics and pragmatics*

Rules or norms do not just lie inertly there; rather, they only exist as rules or norms of *acting*. The social view of language, outlined above, suggests that the meaning of a linguistic expression is identical with its (conventionalized) *use*: "Look at the sentence as an instrument, and at its sense as its employment" (Wittgenstein 1958: §421). Here as elsewhere, the form of an instrument is a *means* to achieve different *goals*. Language use, i.e. speaking, is part of the same general means – ends hierarchy as are all human actions and activities.

Both meaning and its study are called *semantics*. More precisely, semantics is that part of (the study of) meaning which deals with meanings of words and sentences at the ('general') level of the conventional linguistic system, and not at the (concrete) level of single acts of speaking. However, the 'actionist' nature of language is present already in semantics. As a semantic entity, a sentence like *I*

*will come to see you at midnight* encodes an act of *asserting*. The acts of requesting and asking are encoded in imperative and interrogative sentences.

Language is not just action, but also interaction. In the case of requests and questions (codified as corresponding imperative and interrogative sentences) this is self-evident because they can only be conceptualized as being directed to someone different from the speaker himself. But the same is also true of assertions, codified as corresponding declarative sentences, as Sibawaihi, the founder of Arab linguistics, was perceptive enough to realize:

> This is how we speak, even if the listener does not ask loud, because what you say follows the extent of the question he might pose if he were to ask you (cf. Itkonen 1991: 155–156).

The same insight was achieved e.g. by Russell (1967 [1940]: 24):

> In adult life, all speech ... is, in intention, in the imperative mood. When it seems to be a mere statement, it should be prefaced by the words 'know that'. We know many thingsand assert only some of them; those that we assert are those that we desire our hearers to know.

*Pragmatics* is that part of the study of meaning which deals with how the general meaning determined by the linguistic system becomes concrete orspecific in single, either real or imaginary acts of speakings. This requires taking *contextual* information into account. In semantics, as noted above, the sentence *I will come to see you at midnight* has just the meaning of an *assertion*. (Which assertion? — this is evident from the lexical content.) In pragmatics, the same sentence (once uttered) becomes — depending on the context — either a *promise* (= Romeo is speaking to Juliette) or a *threat* (= a vampire is speaking to his future victim). This, in my view, is the relationship between semantics and pragmatics in a nutshell. It coincides with de Saussure's (1962 [1916]) classic distinction *langue* vs. *parole* (see Section 3.1 below).

It may seem natural to assume that pragmatics, concentrating on individual performance, pertains to psychology. In my view, however, pragmatics too is of *social* character. First, the performance is not individual but inter-individual, i.e. it necessarily takes place between speaker and hearer. Second, this inter-individual performance is publicly observable, and derives its identity from being (commonly) understood as a joint result of convention and context; just think of the Romeo vs. vampire contrast (cf. Leech 1983, Verschueren 1999). The truth of this statement remains unaffected by the fact that *psychological* explanations may of course be provided for any type of behavior (including linguistic interaction).

In sum, semantics is the study of context-independent meaning whereas pragmatics is the study of context-dependent meaning. This 'context-independent vs. context-dependent' distinction was captured by Paul (1880 [1975]: Chapter 4) by means of his terminological dichotomy *usuelle* vs. *okkasionelle Bedeutung* (= 'usual vs. occasional meaning'). Sometimes it has been claimed that the

(inter)actionist nature of language becomes evident only in pragmatics. We have just seen that such a view is mistaken. At the level of semantics any sentence encodes a 'frozen action', and it is the task of pragmatics to 'melt' it (cf. Itkonen 1983: 152–164). It is also clear that the 'acts' of referring and predicating belong already to semantics, and not just to (discourse) pragmatics.

The relation between semantics and pragmatics is 'dynamic' in the sense that when context-dependent meanings recur, they may *conventionalize* and thus become part of the linguistic system. This kind of 'ascent' from speech (*parole*, *okkasionelle Bedeutung*) to language (*langue*, *usuelle Bedeutung*) is in general characteristic of *language change* (cf. Section 3.1 below).

Having defended the social view of meaning (and of language in general), I may add a few words on why I find its opposite, i.e. the psychologist view of meaning, less convincing. To be sure, 'psychologism' may mean many different things, and in what follows I shall briefly deal only with one version of this doctrine.

It is not uncommon to see meaning equated either with (unconscious) *schema* or with (conscious) *mental image*. First, let us assume that meanings are schemas. These are hypothetical entities: we do not know what they are, but only presume what they might be; and they may even be non-existent. (Implausible as this may sound, it is certainly possible.) In contrast, we do know the meanings of words like *midnight* and of sentences like *I will come to see you at midnight*; it makes no sense at all to assume that they are non-existent. Therefore meanings cannot be schemas. — It needs to be added immediately that we know the meanings of words and sentences only at the *pre-theoretical* level, i.e. we know them merely as the data. We do *not* know how they should be theoretically analyzed (cf. Section 2.4).

Second, let us assume that meanings are mental images. These are *subjective* or vary from one person to the next whereas meanings are *intersubjective*. (For instance, the sentence *I will come to see you at midnight* has only one meaning in the English language, not as many meanings as there are speakers of English.) Moreover, mental images may be non-existent. Even for a single speaker, there seems to be no mental image (or set of mental images) systematically and reliably connected e.g. with the word *if*. But if we accept the equation 'meaning = use', the meaning of *if* ceases to be a problem. It is enough to state (or list) its different uses: the transition from cause to effect (= *If it is raining during the night, the streets will be wet in the morning*) or from effect to cause (= *If the streets are wet in the morning, it has been raining during the night*), and so on. (But notice again that knowing the different uses of *if* does not entail knowing how they should be theoretically described).

In addition to these specific arguments against viewing meanings as schemas or mental images, we should heed the more general or philosophical admonition voiced by Wittgenstein (cf. above): Pictures or images (including

'schemas') are *never* enough. They must always come equipped with instructions about how they are meant to be interpreted.

When the psychologistic conception of meaning amounts to equating meaning not with any specific mental image, but with subjective experience in general, it seems to be based on the following type of fallacy:

> In order to get clear about the meaning of the word *think* we watch ourselves while we think; what we observe will be what the word means! — But this concept is not used like that. (It would be as if without knowing how to play chess, I were to try and make out what the word *mate* meant by close observation of the last move of some game of chess.) (Wittgenstein 1958: §316).

Accepting the equation 'meaning = use' has a both clarifying and liberating effect. An enormous amount of time and energy has been wasted on trying to solve the problem of 'how meaning exists'. But no one is — or need be — worried about how the use of a hammer or of a computer 'exists'.

## 2. **The ontology of the social**

### *2.1 Physical and social reality*

The ontology of social entities is fundamentally different from the ontology of physical entities:

> There existed electrical storms and thunder long before there were human beings to form concepts of them or to establish that there was any connection between them. But it does not make sense to suppose that human beings might have been issuing commands and obeying them before they came to form the concept of command and obedience (Winch 1958: 125).

The concept of 'command' is such as to be accessible to *consciousness*: commands exist only insofar as they are recognizable as, or known to be, what they are. This type of knowledge must be *shared* by all those who issue commands (and either obey or disobey them). In what follows it will be called *common knowledge*. This provides us with a preliminary definition of 'social': Social entities (unlike physical entities) exist if, and only if, they are commonly known to exist. For instance, money ceases to exist, i.e. it is just pieces of metal and paper, as soon as people no longer know that it exists (qua money).

This definition has some interesting consequences. Because a language like English exists if, and only if, it is commonly known to exist, it follows, among other things, that the correctness of correct sentences is a *social fact*, as elucidated by the following equivalence:

(1) The sentence *John is easy to please* is a correct sentence (of English) iff the sentence *John is easy to please* is commonly known to be a correct sentence.

The formulation (1) is equivalent to the following formulation:

(2) The sentence '*John is easy to please* is a correct sentence' is true iff the sentence '*John is easy to please* is a correct sentence' is commonly known to be true.

The sentence (2) instantiates the Tarskian 'T-sentence', which is of the following general form (cf. Itkonen 1983: 112):

(3) X is true iff p

Here 'p' represents the truth condition of *X*. According to the received view, the *truth value* and the *truth condition* are two different things: we always know the truth condition of *X*, i.e. 'p', and we analyze it in a step-wise fashion, but this happens independently of whether we know *X* to be true or false. As far as physical facts are concerned, it is indeed the case that while we do know the truth condition of *X*, we do *not* know the truth value of *X*. Now, the example (2) refutes the received view as applied to *social facts*, because it shows that, in this crucial domain, it is impossible to know the truth condition of *X* without knowing the truth value of *X* (for discussion, cf. Itkonen 1983: 129–135). Thus, at the level of social facts, the T-sentence has the following form:

(4) X is true iff X is (commonly) known to be true

A declarative sentence *X* is used to make a statement (or assertion). In logical semantics, the truth-condition of *X* is equated with the *meaning* of *X*. This view is too restrictive, but it is certainly the case that knowing the truth condition of *X* is *part* of knowing the meaning of *X*. We have just seen that, in connection with social facts, knowing the truth condition (and, more generally, the meaning) of *X* entails knowing the truth value of *X*. But why should we think, in the first place, that we know the meanings of the words and sentences that we utter? Wittgenstein (1969: §370) suggests the answer: "I should stand before an abyss if I wanted so much as to try doubting their meanings..."

*2.2 The nature of common knowledge*

What does it mean to say that a social entity like the English language is an object of common knowledge? One way to answer this question, due to Lewis (1969), is to say that X is an object of common knowledge if, and only if, the three

conditions given in (5) are true of X and of (practically) any two members of a community (where both 'A' and 'B' stand for each of the two):

(5) A knows-1 X
A knows-2 that B knows-1 X
A knows-3 that B knows-2 that A knows-1 X

As abstruse as such a formulation may seem at first, it is quite easy to show that three-level knowledge of this kind necessarily occurs in all institutional encounters. Suppose I want to cash a check in a bank. The only reason why, when approaching the counter, I do not make soothing gestures or shout "I know what to do, you don't have to tell me!", is that I possess the relevant three-level knowledge: Not only do I know-1 what to do; and not only do I know-2 that the teller knows-1 what to do; but I also know-3 that the teller knows-2 that I know-1 what to do. This type of 'third-level mentality' is also discussed and exemplified by Zlatev (this volume).

From the logical point of view, there is no way to stop the infinite regress of different knowledge-levels (= 'I know that he knows that I know that he knows...'). From the practical point of view, however, this is not a problem. People do not generally go beyond three- or four-level knowledge. Some people are able to do this; but nobody masters e.g. ten-level knowledge.

The explication of 'social' in terms of many-level *knowledge* has sometimes been regarded as entailing some sort of philosophical idealism. Our example of check-cashing behavior should dispel this misunderstanding. The relevant common knowledge is *'embodied'* not just in people's behavior, but also in such physical artefacts as the bank building, its furniture, the clerks' implements, and so on. Sinha (1988) rightly emphasizes the importance of taking into account the *material grounding* of institutions (including language).

Our example is apt to illuminate another often-misunderstood aspect of common knowledge. My attitude vis-à-vis the bank teller is not invalidated if it later turns out that at the moment of our mutual encounter he happened, for instance, to be either unconscious or suffering from an attack of insanity, which means that he did *not*, as a matter of psychological fact, possess the requisite three-level knowledge about me. A's three-level knowledge about B is not about what B knows in fact, but what A is *entitled to expect* B to know: Given the surroundings, I was entitled to expect that the bank teller whom I was approaching knew his business, i.e. had the requisite three-level knowledge about me. Hence, common knowledge turns out to contain a *normative* element. It is a 'rational reconstruction' of sociality, not a psychological description of what actually goes on in people's heads in each and every case:

For in most social situations, if not in all, there is an element of rationality. ...I refer to the

possibility of adopting, in the social sciences, what may be called the method of logical or rational construction, or perhaps the 'zero method'. ...The 'zero method' of constructing rational models is not a psychological but rather a logical method (Popper 1957: 140–141, 158; for discussion, see Itkonen 2003b: 131–135).

The notion of common knowledge has been generalized, and *conventionalized* (cf. Section 3.1), out of single instances of (non-normative) third-level mentality, as described by Zlatev (this volume). But, as it is now, it does contain an ineluctable normative element. The original version of common knowledge given in Lewis (1969) can be criticized for having ignored precisely this fact (cf. Itkonen 1978: 182–186). In sum, the social world (explicated by means of the notion of common knowledge) is permeated by normativity considerations through and through:

It is perhaps the basic insight of Winch (1958) that we need criteria, whose use is governed by rules [= norms], to identify entities as same or different, and that as regards social entities, such criteria are *internal* to them (Itkonen 1978: 185).

Clark (1996) too considers a language as an object of common knowledge, and he claims (pp. 75–77), more precisely, that a language qua commonly known is a set of conventions. This agrees perfectly with my view (even if I prefer the term 'norm'). The conventions include those for 'lexical entries' and those for 'grammatical rules', i.e. norms for pairing (morphemic and lexical) forms with meanings and those for combining meaningful forms into phrases and sentences, as I would say.

Common knowledge (like knowledge in general) must have a *basis*. In the simplest case, the common knowledge of a fact is based on its intersubjectively observable existence. For instance, the common knowledge that it is raining right now is based on the fact that (as everybody can see) it *is* raining right now. But remember that a physical fact, unlike a social fact, can exist, and typically does exist, even if it is not commonly known to exist.

What is the basis for *linguistic* common knowledge, e.g. for (2) in Section 2.1? It cannot be pinpointed as easily as it can in the case of commonly known physical facts. It is not a particular happening, like someone uttering *John is easy to please* and no-one protesting its incorrectness. (To be sure, linguistic common knowledge must not — in general — conflict with such particular happenings.) The basis for common knowledge about the (in)correctness of sentences is 'diffuse', in the sense that it is constituted just by general facts about coming to master a language and by the concomitant common knowledge about those facts. In this respect linguistic common knowledge is just one instantiation of institutional common knowledge in general. The most important difference vis-à-vis common knowledge about physical facts resides in that the basis for linguistic common knowledge, though undeniably existent, cannot be used to strengthen or

justify that which it is a basis for:

> And here the strange thing is that when I am quite certain of how the words are used, have no doubt about it, I can still give no grounds for my way of going on. If I tried I could give a thousand, but none as certain as the very thing they were supposed to be ground for (Wittgenstein 1969: §§ 306–307).

### *2.3 A solution to the controversy between individualism and collectivism*

The definition of social ontology given in Section 2.2 dissolves rather than solves a long-standing controversy within the philosophy of the social sciences. One side has argued that there is an ontological level of social institutions distinct from the level of individual persons. The other side has argued that there are *nothing but* individual persons (cf. O'Neill 1973). Now we can see that they are both right. Indeed, there are nothing but individual persons; but what we have is not just an aggregate of individual persons endowed with arbitrary mental states and distributed in a random order; rather, we have individual persons endowed with quite specific mental attributes (namely many-level states of knowledge) placed in a quite definite structure or pattern (namely that characteristic of common knowledge). It is this structure which constitutes the ontological level of social phenomena.

As an analogy, consider the distinction between a single line and a net. On the one hand, it can be argued that a net consists of *nothing but* lines, which means that the line is ontologically primary vis-à-vis the net. On the other hand, the net is not just a random heap of lines, but a quite specific structure or *pattern* of lines. When the lines constitute a net-like structure, then — and only then — there is this all-important difference that it is possible to catch fish with a net, but not with a line. This difference is important enough to be called 'ontological'; and it shows how increasing complexity makes a new ontological level 'emerge' out of an ontologically simpler level. It could also be argued that in (dis)solving the controversy between individualism and collectivism, we *eo ipso* show that the contrast between psychological and social, which was taken for granted in Section 1, is more apparent than real. In so doing, we have been forced to *revise* the meanings of these two words, i.e. 'psychological' and 'social', to some extent

The preceding discussion suggests that the metaphor of 'social *network*' should be taken seriously. The same analogy may also illustrate the distinction between (subjective) *intuition* and (intersubjective) *norm*, which may at first seem a little puzzling.

Institutions consist of norms. Norms are learned on the basis of observation, but once they are known, they can no longer be just a matter of observation because they are made use of to judge whether an observed (or imagined) action is correct or not:

The correctness of a performance is not among its perceptual characteristics; it cannot be, since it is a relation between the performance and an adopted rule [= norm] — a relation which is more fully expressed by the statement that the performance conforms to the adopted rule (Körner 1960: 117).

The subjective (non-observational) knowledge of norms is called *intuition*. It is a general truth, labelled 'Hume's guillotine', that knowledge of norms (i.e. of what *ought* to be done) cannot be reduced to observation (of what *is* done).

In the definition of common knowledge, it is the first level, i.e. 'A knows-1 X', which corresponds to that standard type of (subjective) linguistic intuition which is used in gathering the data that constitutes the basis for grammar-writing: 'A knows that *y* is a correct sentence'. The second and third levels are also of 'intuitional' character; but more importantly, they bring out the *interactional* nature of language or of social facts in general. Moreover, there is also theoretical understanding about the three-level knowledge as a whole: Although I am just one knot in the social network, i.e. a single person qua member of an institution, whose knowledge and action constitute just a small contribution to its existence, it is nevertheless possible for me to reflect on the institution as a whole.

The 'social world', understood as an object of common knowledge, is co-extensional with Popper's (1972) 'world-3', though without the latter's Platonist overtones. The ineluctably interactional nature of all social facts was beautifully captured by Marx &Engels (1973 [1846]: 37):

Es zeigt sich hier, dass die Individuen allerdings *einander* machen, physisch und geistig, aber nicht *sich* machen. (= So we see that in a physical and spiritual sense individuals make *each other*, but do not make *themselves*.)

### *2.4 Normativity in language*

The fundamental distinction between linguistics and any genuine natural science consists in the fact that the subject matter of the former is inherently normative whereas the subject matter of the latter is inherently non-normative. Now the notion of normativity needs to be explicated more narrowly.

First of all, we have to establish the distinction between a *rule-sentence* such as (6), which describes a rule (or norm), and an *empirical hypothesis* such as (7), which describes an (assumed) regularity.

(6) In English, the definite article (i.e. *the*) precedes the noun (e.g. *man*)
(7) All ravens are black.

The difference between (6) and (7) consists in the fact that (6) can be (and in fact has been) falsified by spatiotemporal occurrences, namely non-black ravens, whereas (6) is not, and cannot be, falsified. The utterance of a sentence (8)

does not falsify (6). Why? — because this sentence is *incorrect*. Nor does the utterance of a sentence like (9) falsify (6). Why? — because this sentence is *correct*. Thus, (6) is unfalsifiable (on the basis of spatiotemporal occurrences).

(8) *Man the came in.
(9) The man came in.

The difference between rule-sentences and empirical hypotheses has been occasionally recognized in the philosophy of the social sciences, e.g. by Ryan (1970), who, to be sure, fails to distinguish between rules (= object of description) and rule-sentences (= description):

> A causal generalization has only one task to fulfil, namely telling us what will and will not happen under particular conditions, irregularities are thus falsifying counter-examples to the causal law. But rules [i.e. rule-sentences] are not falsifiable in any simple way — except of course that it may be false to say that there is a rule — and breaches of a rule are *errors* on the part of those whose behavior is governed by it (p. 141; emphasis added).

In general, however, the distinction at issue has remained in some sort of methodological limbo. On the one hand, one may be willing to admit that perhaps — just perhaps — there may indeed exist something that resembles this distinction. On the other hand, one refuses to draw any methodological consequences from the (possible) existence of this distinction.

What is at issue here is the *normativity* of language: sentences are normative (i.e. correct or incorrect) entities whereas birds are not (or, at least, not in the same sense as sentences are). The normativity of language is ignored in traditional philosophy of language, as shown by the fact that the distinction between sentences and (e.g.) birds is ignored. At the face of it, this is a curious fact, because philosophy of language is brimming with talk about rules of language. In practice, however, no *examples* of these rules are ever given. Because the discussion is carried out at such a high level of generality, the distinction between sentences and (e.g.) birds is destined to remain hidden. — Among philosophers of language, to be sure, there are some laudable exceptions, for instance Cavell (1971a [1958], 1971b [1962]).

In reality, the meanings of words are all based on corresponding *rules*: there are rules which determine that *three* designates a number, i.e. 3, and not a plant, whereas *tree* designates a plant and not a number; and so on for all words of all languages. These rules attach meanings to forms. And then there are rules that determine how meaningful forms have to be combined. One rule of this kind is described by our rule-sentence A. Other such rules deal with facts of government (= 'rection') and agreement. It is *correct* to say *I confided in him* and *incorrect* to say *I confided from him*; it is *correct* to say *I am upset* and *incorrect* to say *You am*

*upset*; and so on. As noted before, Clark (1996) assumes the existence of two corresponding types of rules. For any rule it is possible to construct a corresponding rule-sentence.

The status of rules may be clarified by the following remarks:

> The problem for the grammarian is to construct a description ... for the *enormous mass of unquestionable data* concerning the linguistic intuition of the native speaker (often himself) (Chomsky 1965: 20; emphasis added).

> Few users of language know much in any systematic way about their language, though obviously they can discover any number of *odd bits of correct information* simply through self-observation (Hockett 1968: 63; emphasis added).

Because of their trivial or pre-theoretical character, rules and corresponding rule-sentences possess no linguistic (or scientific) interest whatever. However, their philosophical (or metascientific) significance is enormous. They show that, contrary to what is the case in the natural sciences, the *basic data* of grammatical description are not particular entities (= single spatio-temporal occurrences), but general entities (= norms) described, in principle, by general *and* unfalsifiable sentences. This insight constitutes the core of 'response-dependency' (mentioned in Sect. 1.1).

The standard reaction to what precedes is to say that if the rules/norms of language are known in an unfalsifiable way, or with certainty, there is nothing left for the grammarian or linguist to do. But consider the case of Panini (c. 400 BC), "the greatest grammarian of all" (Dixon 2002: 145). At the pre-theoretical level, his contemporaries knew Sanskrit just as well as he did. But only he was able to construct the grammar that was to bear his name. Thus, once the data are in, *everything* still remains to be done. Similarly, Chomsky and Hockett clearly imply that there is a job for them to do, whatever odd bits of correct and indubitable information the average speaker may possess about his language.

The same point can be made by briefly returning to the notion of truth condition. As Wittgenstein so eloquently put it, we stand before an abyss if we start to doubt whether or not we know the meanings of the words and sentences that we use. But of course we know them only at the pre-theoretical level. We know that *John is easy to please* is a correct English sentence (unlike e.g. **John is easy from please*) and that it means the opposite of *John is difficult to please*, but we do *not* know the best theoretical description of this (or any other) sentence. Any theoretical description is falsifiable by definition. But falsification in grammatical description is not what it is in the natural sciences.

There are many other 'standard objections' against the distinction between rule-sentence (= A) and empirical hypothesis (= B), for instance:

- "If English were different, A would be falsified."
- "In English (as it is now) A is verified and any other formulation of the same facts is falsified."
- "The definite article does not (always) precede the noun (just think of *Ivan The Terrible*)."
- "Maybe A is not falsifiable by simple observation, but neither are scientific theories."
- "The terms 'definite article' and 'noun' are theoretical, not pre-theoretical."
- "A and B are formulated in dissimilar ways."
- "Not all rules of English are of the same type as the one described by A."
- "The existence of the rule described by A is a contingent and not a necessary fact."
- "A is not an analytical sentence."
- "English has also statistical and experimental aspects not captured by A-type sentences."

Such and similar objections have been brought together and answered in Itkonen (2003b: Chaps 3, 6, 7); see also Section 3 below.

It should also be pointed out that the mere existence of the normativity of language is enough to refute all varieties of physicalism (or 'naturalism'), i.e. of the view that physical data is all there is. If you *argue* for this view, you must do so in the *language* of physics (and/or philosophy); and the language you use is not physical (or 'naturalistic'), but normative.

*2.5 Correctness vs. rationality*

In typical linguistic behavior, rational actions are performed by uttering correct sentences. It is quite possible, however, to perform irrational actions by uttering correct sentences, and to perform rational actions by uttering incorrect sentences, which shows that the dimensions of correctness and rationality are independent from each other.

Since the use of language exemplifies the general means – ends hierarchy, as noted in Section 1.2, it is amenable to so-called *rational explanation*, which is a general explanatory model for human (and even animal) behavior:

> To explain an action as an action is to show that it is rational. This involves showing that on the basis of the goals and beliefs of the person concerned the action was the means he believed to be the most likely to achieve the goal (Newton-Smith 1981: 241).

Even *irrational* behavior can be explained, if at all, only by means of

rational explanation, namely by exposing the *reason* why it was performed. This involves coming to understand how behavior that *is* irrational in fact came to *seem* rational to the agent. The 'transition' from goals to means followed by the carrying-out of the means, as codified in rational explanation, can be seen as the *causal force* that brings about linguistic behavior investigated in such distinct linguistic subdisciplines as psycholinguistics, sociolinguistics, and diachronic linguistics (cf. Itkonen 1983).

> Using language must consist of the continuous making of linguistic *choices*, consciously or unconsciously, for language-internal (i.e. structural) and/or language-external *reasons* (Verschueren 1999: 55–56; emphasis added).

This innocuous-looking statement, once its implications are spelled out, justifies the use of rational explanation.

## 3. Normativity and beyond: language change, language psychology and typology

### *3.1 Language change: The need for statistics*

Language change entails that old norms (or rules) are replaced by new ones. Comparative Indo-European linguistics started with the idea of *grammaticalization*. Thus, Franz Bopp claimed in 1816 that, for instance, the endings of Sanskrit verbs had originally been full personal pronouns (cf. Arens 1969: 177). To give another example, let us consider the Modern French constructions *venir de + INF* and *aller INF*. Originally these had the concrete local meanings 'come from INF' and 'go INF'. Then in some contexts these constructions were *reanalyzed* as having also the temporal meanings 'recent past' and 'near future'. First, these meanings were more or less accidental or *pragmatic*; but later they became conventionalized or *semantic*. (As noted in Section 1.2, this 'pragmatic vs. semantic' distinction is just a reformulation of Paul's (1975 [1880]) distinction between *okkasionelle* vs. *usuelle Bedeutung*.) That new conventions or norms had emerged, was evident as soon as the temporal meanings were *extended* to such contexts where the old concrete and non-temporal meanings are impossible, as shown in (10) and (11).

(10) *Il vient de mourir* ('he has just died' < 'he comes from dying')
(11) *Il va s'éveiller* ('he will wake up' < 'he goes wake up').

The mechanism of grammaticalization (= reanalysis-cum-extension) is discussed e.g. in Itkonen (2002). It is a curious fact that while in theoretical linguistics much attention has been devoted to the notion of conventionalization,

the logically primary notion of convention (or normativity) has remained practically unknown.

The (typical) linguist takes the existence of language for granted. He is not competent by training to answer the phylogenetic question concerning the origin of language. Nor is it his business to reconstruct the process through which norms may have emerged out of an attempt to coordinate originally non-normative actions (cf. Lewis 1969). This does not mean, however, that these are not worthwhile questions to be asked in an interdisciplinary framework.

Traditionally, grammarians have been relying on self-invented example sentences, which means that traditional *synchronic* linguistics has been based on *intuitional* data (for extensive documentation, see Itkonen 1991). The use of intuitional data unites such otherwise dissimilar approaches as generativism (= Chomsky 1965, Jackendoff 1994), cognitive linguistics (= Lakoff 1987, Langacker 1987), and construction grammar (= Goldberg 1995, Croft 2001). The reliance on intuitional data is fully justified in so-called *clear cases* (exclusively focused upon by the six linguists just mentioned), but elsewhere one has to resort to observation of actually occurring utterances, which entails the use of *statistics*.

Norms of language may be more or less binding, i.e. they may determine the correctness of expressions or sentences either in a discrete ('either – or') way or in a non-discrete ('more-or-less') way. In most languages, for instance, the norms of word order are non-discrete while the norms of affixal morphology are discrete. The norms of word meaning are open, in the sense that there is a discrete core surrounded by a non-discrete periphery: "It is only in the normal cases that the use of a word is clearly prescribed;..." (Wittgenstein 1958: §142).

Even when the norms are discrete, the (normative) behavior they subsume is non-discrete, which is another way of saying that they may be broken (either deliberately or inadvertently). A much discussed example is the *t/d* deletion in today's English (cf. Hudson 1997). The (discrete) norms determine the phonological form of the noun *mist* ('fog'), the past tense *left* of the verb 'to leave', and the past tense *missed* of the verb 'to miss'. But in actual practice, the word-final *t/d* may or may not be present, and in these three cases it is typically retained in the following proportion: 50% – 65% – 80%. There is the experience of this statistical pattern (based on observation), in addition to the (intuitive) knowledge of the above-mentioned discrete rules. This duality can be captured by assuming that what a discrete norm determines is a *prototype*: while a prototype is defined by its 'typical' features, any of these may be overridden in exceptional cases. The important thing is that this duality must not be explained away. In particular, it would be wrong to try to reduce the discrete norm to the corresponding non-discrete and statistical behavior. This follows from the fact, mentioned in Section 2.3, that 'ought' cannot be reduced to 'is'.

When the percentage of the norm-following behavior drops below 50%, at the latest, we are witnessing a *diachronic* process which turns a discrete norm into

a non-discrete one and, in general, ultimately leads to its disappearance. This amounts to a change of the prototype, which in turn equals a *language change*. This can be a lengthy process. For instance, in one hundred years the correct pronunciation of today's *mist* may actually be *[mis]*. To give a less speculative example, it took some 300 years (i.e. between 1450 and 1750) for the construction exemplified by (12) to be replaced by the construction exemplified by (13) as part of the emergence of the auxiliary system of Modern English.

(12) Saw he the dragon?
(13) Did he see the dragon?

First, the latter structure was totally incorrect, and in the end it came to be totally correct. In between, there was a gradual shift that can be described only in *statistical* terms (cf. Hudson 1997). In other words, language change is a prime example of *less-than-clear cases*.

It is easy to see that Saussure's terminological distinction between *langue* and *parole* captures the following dichotomy: on the one hand, language as a system of norms accessible to conscious intuition; on the other, actual spatio-temporally specifiable linguistic behavior that is accessible to observation.

### *3.2 Language and the psychology of language: The need for experimentation*

"La langue est une institution sociale" (Saussure 1962 [1916]: 33). It is a general fact that an institution or, more generally, any rule-system S can be described or formalized in many different ways. This means that different people may view S from different perspectives and with different descriptive goals in mind. Thus, there is no a priori reason to assume that the description of S must aim at capturing the way that S has been internalized by those who have learned it. For instance, it is possible to describe S so as to achieve either a maximal degree of operational efficiency or a maximal degree of logical simplicity. The types of descriptions of S that result from adopting either one of these two perspectives will differ from each other, just as they will both differ from the type of description of S that sets the psychology of the users of S as its goal:

> But what would that grand success [of sequence-extrapolating algorithms] teach us about human perception, pattern recognition, theory formation, theory revision, and esthetics? Nothing — nothing at all.
>
> This ... brings out the vastness of the gulf that can separate different research projects that on the surface seem to belong to the same field. ... Today's wonderfully powerful chess programs, for instance, have not taught us anything about general intelligence — not even about the intelligence of a human chessplayer!
>
> Well, I take it back. Computer programs *have* taught us something about how

human chessplayers play — namely, how they do *not* play. And much the same can be said for the vast majority of artificial-intelligence programs (Hofstadter 1995: 52–53).

This is a very clear formulation of the fact that there is a difference between a description of S, or D1, and a description of the *psychology* of S (= P-S), or D2. Thus, D1 and D2 refer to, and describe, two distinct entities, namely S and P-S. The understanding of this distinction has been made needlessly difficult by ambiguous terminology. On the one hand, P-S is often called 'knowledge of S'. On the other, S is — by definition — commonly known. This creates the wrong impression that there is no difference between S and P-S nor, consequently, between D1 and D2.

For the sake of clarification, consider the following analogy. If I describe the moon as I see it with the aid of a telescope, it is still the moon that I describe, and not my vision (enhanced by the telescope). If I genuinely wish to concentrate on my vision, and not on the moon, then I have moved from astronomy to the psychology of vision. Exactly the same remarks apply to the distinction between D1 and D2, as Hofstadter so well demonstrates. It is only D2 which aims at *psychological reality* whereas D1 has other desiderata (e.g. efficiency or simplicity).

Once you have grasped this distinction, you realize that it applies practically everywhere. For instance, there is a difference between geometry and the perception of geometrical figures and shapes (cf. Itkonen 1983: 1–3). In just the same way, there is a difference between formal logic and psychology of logic (cf. Itkonen 2003a: Chap. XV). In linguistics, the matters may at first seem less clear. Therefore it is good immediately to point out that there are quite uncontroversial cases of *non*-psychological grammatical descriptions. For instance, it is a fact, pointed out by Paul Kiparsky (p.c.), that Panini's grammar does *not* strive after psychological reality. Similarly, in arguing against the view that linguistics is psychology, Katz (1981) operates with the concept of 'optimal grammar':

> [There should be no] constraints that impose a ceiling on the abstractness of grammars by tying them down to one or another particular [i.e. physical or psychological] reality (p. 52).
>
> A grammar *G* is an *optimal grammar* for the language *L*, if ...*G* ...implies every true evidence statement about *L* ...and there is no grammar *simpler* than *G*... (p. 67; emphasis added).
>
> [O]n the most natural definition, an 'optimal grammar' is a system of rules that predicts each grammatical property and relation of every sentence in the language and for which there is no simpler (or otherwise methodologically better) such predictively successful theory (Katz 1985: 201; original emphasis deleted).

However, Katz's references to 'optimal grammar' remain rather unconvincing, because he is unable to *exemplify* this concept. Therefore it is

important to emphasize that, within the 'world history' of linguistics, this concept has already been exemplified rather well, namely by Panini's grammar:

> [Panini's grammar] is the most comprehensive generative grammar written so far (Kiparsky 1979: 18). Modern linguistics acknowledges [Panini's grammar] as the most complete generative grammar of any language yet written, and continues to adopt technical ideas from it (Kiparsky 1993: 2912).

The same laudatory view of Panini's grammar has been both documented and argued for in Itkonen (1991: Chap. 2, esp. pp. 68–70). In the present context it is important to understand that, in addition to being the best *generative* grammar, Panini's grammar is — by Kiparsky's own admission (cf. above) — also a *non-psychological* grammar, which means that it is indeed a serious candidate for being the Katz-type 'optimal grammar'.

The notion of non-psychological or 'autonomous' linguistics has been analyzed in Itkonen (1978) and Kac (1992). Katz (1981), (1985) gives it a Platonist interpretation, but there is really no reason to do so:

> The properties Katz assigns to abstract objects appear all to be possessed by the kind of conventions of mutual knowledge that Esa Itkonen argues are constitutive of linguistic rules (Itkonen 1978; not cited in Katz 1981) (Pateman 1987: 52).

While language is identical with a system of (social) norms, psychology of language is identical with the structures and processes involved in speech understanding and production as well as in the mental storage of linguistic units. In Itkonen (1983) this distinction was conceptualized as holding between (social) norms and (individual-psychological) internalizations-of-norms. It is in connection with the latter that the need for *experimentation* arises. This can be illustrated by means of what is probably the most famous example in recent decades.

The 'standard theory' of generative linguistics, as expounded in Chomsky (1965), made use of a descriptive apparatus consisting of transformations that convert deep structures into surface structures. This is one possible method of presenting intuition-based data in a systematic way; indeed, it was already used by Apollonius Dyscolus, who wrote the oldest extant syntactic treatise of the Western tradition (cf. Itkonen 1991: 206–211). But is it also psychologically adequate? And how can this question be answered, in the first place?

Experimentation provides the answer. If transformations are psychologically real processes, they must take time to be performed. Hence, the hypothesis is that there are longer reaction times connected with producing and/or understanding sentences that involve more (rather than less) transformations. Experimental data give this verdict: "[T]he hypothesis that the operations that the subjects performed were grammatical transformations is actually disconfirmed by the data" (Fodor et al. 1974: 241).

That it is perfectly legitimate to use transformations in grammatical description (= 'autonomous linguistics') in spite of their psychological *non*-reality, shows that, in Hofstadter's (1995) words, there is a "gulf" that separates intuition-based autonomous linguistics from experimental psycholinguistics. More precisely, the data of the former is of *pre-experimental* character; it is a *precondition* for the data of the latter: "One cannot make experiments if there are not some things that one does not doubt" (Wittgenstein 1969: §337).

The existence of pre-experimental linguistic knowledge has occasionally been acknowledged: "It is pointless to run an experiment which shows that if something is a pencil, appropriately motivated English speakers will call it 'pencil'. Anyone who knows English knows that already" (Fodor et al. 1974: 399–400). This type of experiment would be "a slightly absurd exercise, with the results a foregone conclusion" (Wason & Johnson-Laird 1972: 78). However, the larger implications have remained unexplored and poorly understood.

The ambiguity of non-psychological vs. psychological study of language is well illustrated by the notion of *analogy*. On the one hand, analogy may be just a convenient descriptive device for presenting the data. On the other, analogy may be meant to capture the actual structure-cum-process that brings linguistic behavior about (cf. Itkonen 2005a).

### *3.3 The nature of typological linguistics*

Up to now we have come across three distinct types of linguistic data, namely intuitional, observational, and (observational-)experimental. The two latter types deal with frequencies of spatio-temporal occurrences and thus require a statistical mode of description. This division of labor between different linguistic subdisciplines was already set forth in Itkonen (1977) and (1980).

What is the status of typological linguistics from the present perspective? An in-depth analysis of the reference grammars of ten more or less 'exotic' languages reveals a general lack of any statistical means of description (cf. Itkonen 2005b). This shows that, once again, we are dealing with intuitional data. In many cases, however, what we have is not the intuition of a (field) linguist, who, while writing his grammar, may still be in the process of learning the language to be described, but the intuition of his informant(s). In other words, we are dealing with *elicitation*. Haiman (1980: xi) gives an eloquent account of this method:

> I will always remember Kamani Kutane for his thought experiments: given a minimally contrasting pair of sentences, he would construct elaborate background stories which would be appropriate for only one of these sentences. Eventually I would understand one of these, and we could move on. It was by means of such continued thought experiments that he was able to make clear to me the meaning of that most mysterious of all Hua forms, the gerund *-gasi*.

As shown by this quotation, and as argued in Itkonen (2004), the study of 'exotic' languages is based on *empathy* as a form of intersubjectivity, or — in Collingwood's (1946: 218) words — our capacity of "rethinking the same thought which created the situation we are investigating, and thus coming to understand this situation". But once we have become aware of empathy in this context, we realize that we have been using it all the time. For instance, we can explain the grammaticalization of the constructions *venir de INF* and *aller INF* in the way we do (cf. Sect. 3.1), only because we *understand* the processes of reanalysis and extension that are involved here; and we understand them, because we can 're-enact' them, i.e. we realize that we could have *done the same thing*. On reflection, this turns out to be an application of the model of rational explanation (cf. Sect. 2.5). — The discovery of mirror neurons seems to have revitalized the notion of empathy, as is shown in detail by Barresi & Moore (this volume).

## 4. The roots of the anti-normative bias in theoretical linguistics

Why has there been such a pronounced inclination to ignore the ineluctably normative character of language? There are many reasons, of which I mention here only two. First, there is sheer intellectual laziness:

> [It is wrong] to consider the salient features of an object as representative of its totality. In this way the evident concreteness of the sound of words leads one to ignore the extent to which *use*, however intangible, is necessary to word-hood (Friedman 1975: 94, emphasis added; discussed in Itkonen 1978: 182–183).

Notice that it is the same, or very similar, fallacy that underlies the entire Cartesian tradition mentioned in Section 1.1. This is the Cartesian argument in outline: "I see, and hence I know, that this thing in front of me is a burning candle; but I do not see anyone else in the room; thus when I know what I know about the thing in front of me, I am alone; therefore my knowledge is not social but subjective; and what is true of my knowledge here and now is true of every type of knowledge." Once this argument has been spelled out, one cannot help marvelling how simple, and simple-minded, it really is.

Second, there is the temptation to replace the (normative) 'correct vs. incorrect' distinction by the (non-normative) 'possible vs. impossible' distinction. Thus, Jackendoff (1994: 49–50) claims that, unlike a sentence like 'Harry thinks Beth is a genius', a sentence like 'Amy nine ate peanuts' is "not a possible sentence of English". However, it is not only the case that this is a *possible* sentence of English. We see with our own eyes that it is also an *actual* sentence of English, namely *incorrect* English. It must be actual because (an exemplification of) it occurs in space and time (cf. Dretske 1974: 24–25, Itkonen 2003b: 142–144).

But why should it be tempting, in the first place, to replace normative by non-normative? — because of the prestige enjoyed by the natural sciences. The data of physics is inherently non-normative. From this, it has been wrongly inferred that the data of linguistics too must be non-normative, come what may.

Is there, then, no normativity in the natural sciences? Of course there is. Just think of *protophysics* which investigates the set of norms for measuring space, time, and mass (cf. Böhme 1976). But protophysics is not physics: "It is one thing to describe methods of measurement, and other to obtain and state results of measurement" (Wittgenstein 1958: §242). As argued in Itkonen (1978: 42–48) and elsewhere, protophysics is in a certain sense a methodological equivalent of autonomous linguistics. Still, this is an imperfect analogy because what protophysics deals with are norms of researchers, not of research objects.

In sum, I have argued in this chapter that normativity is indispensable for the existence of language, and that it has been — often without self-awareness — pivotal for linguistics from its very dawn. To remain blind to this obvious fact, a strong bias has indeed been needed.

*I wish to thank Jordan Zlatev for his help in editing the first draft of this chapter.

REFERENCES

Arens, H. 1969. *Sprachwissenschaft*: *Der Gang ihrer Entwicklung von der Antike bis zur Gegenwart*, Band I. Frankfurt a/M: Athenäum.

Blackburn, S. 1984. *Spreading the Word.* Oxford: Oxford University Press.

Böhme, G. (ed.) 1976. *Protophysik*. Frankfurt a/M: Suhrkamp.

Cavell, S. 1971a [1958]. "Must we mean what we say?" In *Philosophy and Linguistics*. C. Lyas (ed.), 131–165. London: Macmillan.

Cavell, S. 1971b [1962]. "The Availability of Wittgenstein's Later Philosophy. In *Philosophy and Linguistics.* C. Lyas (ed.) 1971, 166–189. London: Macmillan.

Chomsky, N. 1965. *Aspects of a Theory of Syntax*. Cambridge, MA: The MIT Press.

Clark, H. 1996. *Using Language*. Cambridge: Cambridge University Press.

Collingwood, R.G. 1946. *The Idea of History*. Oxford: Clarendon Press.

Croft, W. 2001. *Radical Construction Grammar.* Oxford: Oxford University Press.

Dretske, F. 1974. "Explanation in Linguistics". In *Explaining Linguistic Phenomena*. D. Cohen (ed.), 21–41.Washington, D.C.: Hemisphere Publishing Company.

Dixon, R.M.W. 2002. *Australian Languages*. Cambridge: Cambridge University Press.

Fodor, J., Bever, T. & Garrett, M. 1974. *Psychology of Language*. Cambridge,

MA: The MIT Press.
Friedman, H.R. 1975. “The Ontic Status of Linguistic Entities”. *Foundations of Language* 13-1, 73–94.
Goldberg, A. 1995. *Constructions.* Chicago: University of Chicago Press.
Haiman, J. 1980. *Hua: A Papuan Language of the Eastern Highlands of New Guinea*. Amsterdam: Benjamins.
Haukioja, J.2000. “Grammaticality, Response-Dependency, and the Ontology of Linguistic Objects”. *Nordic Journal of Linguistics* 23, 3–25.
Heil, J. 1992. *The Nature of True Minds.* Cambridge: Cambridge University Press.
Hockett, C. F. 1968. *The State of the Art.* The Hague: Mouton.
Hofstadter, D. 1995. *Fluid Concepts and Creative Analogies*. Penguin Books.
Hudson, R. 1997. “Inherent Variability and Linguistic Theory”. *Cognitive Linguistics* 8-1, 73–108.
Itkonen, E. 1977. “The Relation between Grammar and Sociolinguistics”. *Forum Linguisticum* I/3, 238–254.
Itkonen, E. 1978. *Grammatical Theory and Metascience*. Amsterdam: Benjamins.
Itkonen, E. 1980. “Qualitative vs. Quantitative Analysis in Linguistics”. In *Evidence and argumentation in linguistics*. T.A. Perry (ed.), 334–366. Berlin: deGruyter.
Itkonen, E. 1983. *Causality in Linguistic Theory*. London: Croom Helm.
Itkonen, E. 1991. *Universal History of Linguistics: India, China, Arabia, Europe.* Amsterdam: Benjamins.
Itkonen, E. 1997. “The Social Ontology of Linguistic Meaning”. *SKY: The Yearbook of the Linguistic Association of Finland*, 49–80.
Itkonen, E. 2002. “Grammaticalization as an Analogue of Hypothetico-Deductive Thinking”. In *New Reflections on Grammaticalization*. I. Wischer & G. Diewald (eds.), 413–422. Amsterdam: Benjamins.
Itkonen, E. 2003a. *Methods of Formalization beside and inside both Autonomous and non-Autonomous Linguistics*. University of Turku: Publications in General Linguistics 6.
Itkonen, E. 2003b. *What is Language? A study in the Philosophy of Linguistics*. University of Turku: Publications in General Linguistics 8.
Itkonen, E. 2004. “Typological Explanation and Iconicity”. *Logos and Language* V/1, 21–33.
Itkonen, E. 2005a. *Analogy as Structure and Process: Approaches in Linguistics, Cognitive Psychology and Philosophy of Science*. Amsterdam: Benjamins.
Itkonen, E. 2005b. *Ten non-European Languages: An Aid to the Typologist*. University of Turku: Publications in General Linguistics 9.
Jackendoff, R. 1994. *Patterns in the Mind*. New York: Basic Books.
Kac, M. 1992. *Grammars and Grammaticality*. Amsterdam: Benjamins.
Katz, J. 1981. *Language and Other Abstract Objects*. Oxford: Blackwell.
Katz, J. 1985. “An outline of Platonist Grammar”. In *The Philosophy of*

*Linguistics*. J. Katz (ed.), 172–203. Oxford: Oxford University Press.
Kenny, A. 1975. *Wittgenstein*. Penguin Books.
Kiparsky, P. 1979. *Panini as a Variationist*. Cambridge, MA: The MIT Press.
Kiparsky, P. 1993. "Paninian linguistics". In *The Encyclopedia of Language and Linguistics*, Vol. 1(6). R.E. Asher (ed.), 2918–2923. Oxford: Pergamon Press.
Körner, S. 1960. *The Philosophy of Mathematics*. London: Hutchinson.
Lakoff, G. 1987. *Women, Fire, and Dangerous Tthings*. Chicago: University of Chicago Press.
Langacker, R. 1987. *Foundations of Cognitive Grammar*, Vol. I: *Theoretical Perspectives*. Stanford, CA: Stanford University Press.
Leech, G. 1983. *Principles of Pragmatics.* London: Longman.
Lewis, D. 1969. *Convention*. Cambridge, MA: Harvard University Press.
Marx, K. & Engels, F. 1973 [1846]. *Die deutsche Ideologie*. Werke, Band 3. Berlin: Dietz Verlag.
Newton-Smith, W.H. 1981. *The Rationality of Science*. London: Routledge.
O'Neill, J. (ed.) 1973. *Modes of Individualism and Collectivism.* London: Heinemann.
Pateman, T. 1987. *Language in Mind and Language in Society.* Oxford: Clarendon Press.
Paul, H. 1975 [1880]. *Prinzipien der Sprachgeschichte*. Tübingen: Niemeyer.
Pettit, P. 1996. *The Common Mind*, 2nd ed. Oxford: Oxford University Press.
Popper, K. 1957. *The Poverty of Historicism*. London: Routledge.
Popper, K. 1972. *Objective Knowledge*. Oxford: Oxford University Press.
Russell, B. 1967 [1940]. *An Inquiry into Meaning and Truth*. Pelican Books.
Saunders, J. T. & Henze, D. F. 1967. *The Private-Language Problem*. New York: Random House.
Saussure, F. de. 1962 [1916]. *Cours de linguistique générale*. Paris: Payot.
Sinha, C. 1988. *Language and Representation: A Socio-Naturalistic Approach to Human Development*. New York: Harvester.
Verschueren, J. 1999. *Understanding Pragmatics*. London: Arnold.
Wason, P.C. & Johnson-Laird, P. 1972. *Psychology of Reasoning*. Cambridge MA: Harvard University Press.
Winch, P. 1958. *The Idea of a Social Science*. London: Routledge.
Wittgenstein, L. 1958. *Philosophical Investigations*, 2nd ed. Oxford: Blackwell.
___. 1969. *On Certainty*. Oxford: Blackwell.
Zlatev, J. 2006. "Language, Embodiment and Mimesis". In *Body, Language and Mind*, Vol. I: *Embodiment*. T. Ziemke, J. Zlatev & R. Frank (eds.). Berlin: Mouton.
___. This volume. "Intersubjectivity, language, and mimesis hierarchy"