Searle's Defence of Internalism

Petr Koťátko

The Academy of Sciences of the Czech Republic, Prague

Abstract: The paper argues in favour of the Searlean internalist way of construing the content of thoughts and communicative acts. As interpreted by the author, it reflects the participation of external factors in the determination of content (stressed by externalists like Burge, Putnam, Kripke and Davidson) but it does so in a way which enables to approach the thinking and communicating individual as a real subject of his acts. The crucial move is to incorporate the individual's reliance on the relevant parameters of his physical or social environment into the construal of the content itself. Then the external factors can be treated as exploited by the individual himself in the articulation of his acts, rather than as intervening into the content through the gaps marked by indexicals, proper names, natural kind terms or concepts borrowed from the communal repertoire. In particular with respect to the last factor the author proposes a maxim of interpretation which he labels as The principle of maximization of subjectivity and regards as a pendant of the well-known Principle of charity.

Keywords: internalism/externalism, propositional content, indexicals, proper names, natural kind terms, division of linguistic labour.

The externalists' claim that the content of our communicative acts or of our beliefs, desires and other attitudes is – at least in some respects – determined by external factors and that these factors can play this role without any internal mediation, so that two individuals in type-identical physical and psychological states can entertain different beliefs, desires etc. and express different propositions (by uttering type-identical sentences). The externalist arguments focus on particular components of our thoughts or communicative acts which are supposed to mark gaps in the internal determination of their contents, and hence function as channels through which these external intervences into the content take place. These components include indexicals, proper names, natu-

ral kind terms and all those general terms which are subject to the division of linguistic and intellectual labour.

John Searle suggests a complete internalist (or, as he calls it, *intentionalist* or *Fregean* – cf. e.g. Searle 1983, 22) construal of the meaning of these terms and of the propositional contents expressed by utterances of sentences including these terms. Moreover, he does so in a way which neither neglects nor obscures but rather fully reflects the ways in which our thoughts and communicative acts are embedded in their external environment, so that the externalist initiative is at least in some respects accommodated within the internalist construal of content, rather than to completely go by the board. This can be easily overlooked when we focus purely on Searle's arguments pointing to controversial presumptions behind the externalists' claims and to mistakes in their inferences.

1   An internalist reply to the externalist initiative

In the cases I am going to discuss here one can recognize a general strategy which I am inclined to call *Searlean*, although I will present it in my own words and with a motivation derived from my own priorities in the dispute with externalists. Within this strategy, the internalist addresses the externalist as follows: you claim that you have discovered external factors which, in the circumstances described by you, play an ineliminable role in the determination of content. How should we understand the way in which these factors do their job? In particular, should we suppose that the thinking or communicating subject *relies on them* in his thoughts or communicative acts? The reply which can be found in the prominent externalist theories is "yes": in fact, this is an essential part of their arguments, or at least of the way in which they attempt to make their position intuitively appealing (examples will follow soon). But then the manoeuvre which I would call the *Searlean move* follows quite naturally. If the subject relies on, or counts with, these external factors as involved in the determination of the content thought or communicated by him, oughtn't we to do justice to this reliance by reflecting it in the construal of the content itself? In particular, oughtn't we to include specification of the ways in which these external factors are supposed to do their job into the construal of the content? Obviously, this specification of the role of external factors should exploit exclusively the devices which are in the possession of the thinking subject - otherwise it could not serve to represent the subject's reliance on

these factors. But once this requirement is fulfilled, the participation of these factors in the determination of content becomes specified within the *internally fixed content* itself (within the intentional content, in John Searle's terminology) and then the contribution of the external factors does not mark any parameter or element of the content which would be left internally underdetermined. But then the externalists fail to give us a reason to postulate any content beyond (or any content broader than) internally fixed content.

In this way we can do full justice to the factors involved in the determination of content to which the externalists appeal. In general, this method of construing the content, when applied in reaction to various versions of externalism, is able to reflect the variety of ways in which our thoughts and communicative acts are anchored in their external circumstances. In fact, what could be a more consequent way of appreciating the relevance of these factors than including a specification of them and of the role they are supposed to play, into the construal of the intentional content of thoughts and communicative acts? But when doing this, we no longer present the role of these factors in the determination of content as an *intervention from the outside*: on the contrary, we present the thinker or speaker *himself* as involving these factors into the articulation of the content of his beliefs, intentions, desires etc. as well as of his assertions, promises, orders etc. In other words, we acknowledge him as a *real subject* of his thoughts and communicative acts.

To sum up, what the internalist responding in the Searlean way to the *externalist* arguments should say, when he wants to be polite, is: many thanks for a valuable innovation of the *internalist* construal of content.

2   Natural kind terms

When defending the construal of intentional content which includes a specification of the manner in which external factors are supposed to contribute to it, I have argued that this way we do justice to the thinker's or speaker's reliance on these factors. And I have pointed out that this subjective reliance is presupposed in the prominent externalist theories themselves. To begin with, let's take Putnam's theory of *natural kind terms*, e.g. his account of the term "water" as introduced into our vocabulary to refer to samples of water identified descriptively or demonstratively plus to anything else with the same essence. When jus-

tifying this account, Putnam argues that it is an essential feature of our attitude to the world that we intend our classification of natural entities to respect their essences (their internal structure responsible for the observable properties) and that we take these essences as making our applications of natural kind terms right or wrong independently on our knowledge of them (cf. e.g. Putnam 1975, 244; Putnam 1981, 46f.). But then we have a good reason to reflect this respect to essences, ascribed by Putnam to human beings, in the construal of the content of communicative acts which include utterances of the term "water" ("lemon", "tiger" etc.), as well as of the content of beliefs, desires etc. expressed in these acts. The meaning of the term "water" can then be specified e.g. by the description "anything that shares essence with what flows in rivers (on Earth), falls in drops from the clouds (on Earth) etc." – in the case that the relevant sample is picked out descriptively. Alternatively, it can be specified by the description "anything that shares essence with *this* stuff" if the relevant sample is picked out demonstratively.

The demonstrative "this", as it appears within this description, might be regarded as marking a gap within this allegedly internalist construal of content – a place through which external factors intervene into the content in a way which is not internally mediated. However, here we can appeal to Searle's internalist account of indexicals (cf. e.g. Searle 1983, 228). The resulting internalist specification of the meaning of the term "water" will then be something like: "anything that shares essence with the stuff which causes *this* experience". The demonstrative "this", as it appears within this new construction, ought not to bother the internalist at all, since it plainly refers to something internal, namely to particular experience of the subject in question (the speaker uttering the term "water").

Thus, if we want to construe the content expressed by an utterance of the sentence "Water is indispensable for life" in a way which is completely internalist (exploits only resources which are in the possession of the thinking subject) and at the same time accommodates the core of Putnam's theory, we get: "The stuff causing this experience and anything else with the same essence is indispensable for life." To sum up, Searle seems to be right when he points out in Searle (1983, 204): "Even supposing Putnam is right about his intuitions, all he has done is substitute one intentional content for another."

Even after this move one can still say, with Putnam, that meaning or intension is "extension involving" (the internal structure of particular

samples of some liquid, independently on what we know about it, enters into the principle determining what is and what is not the right application of the term "water"). However, the role of this *external* factor is fully determined in the internally fixed content of the communicative act performed in the utterance of a sentence including the term "water", and hence it is determined *internally*. It would be absurd to object: in your construal of the content you have replaced the external factor to which the externalists appeal, namely the internal structure of natural entities (e.g. the chemical composition of some stuff), by something else, namely by its description. The description "the essence of this stuff" serves precisely to identify the essence of the stuff in question, just like the description "Brown's murderer" serves to pick out a particular person (cf. Searle's polemics with Putnam in Searle 1983, 205f.).

3   Proper names

Similarly, *Kripke's theory of proper names* not only links our utterances of proper names to individuals as their referents through the chain of uses of the name, anchored in the initial baptism (cf. e.g. Kripke 1972, 91f.). It appeals to the *intention* of any user of the name to join the chain – more specifically, to refer, when uttering the name, to the same individual as his predecessor in the chain (cf. Kripke 1972, 96f., 162). This is an essential part of the theory, since this is supposed to connect one segment of the chain with another, and so to guarantee the continuity of the chain and its ability to link the name to its referent. What could then be more natural than to reflect this speaker's respect to the (Kripkean) referential mechanism in the construal of the content expressed by his utterance of a sentence including a name? The intentional content supposed to do the identificatory work in a case in which the speaker utters the name 'Jan Novák' with the plain intention to refer to the same person as the speaker from whom he got the name, without having any independent identifying knowledge, can be specified precisely by means of the description "the man referred to by the speaker from whom I got the name 'Jan Novák'".[1]

---

[1]   If the speaker possesses some non-parasitic identifying knowledge, it is natural to suppose that it has the form of a *cluster* of identifying descriptions – details see in Searle (1958); Searle (1969, Ch. 7); cf. also Zsofia Zvolenszky's paper in this volume.

This is simply a way of reflecting (one of the possible forms of) the parasitic uses of names (cf. e.g. Strawson 1959, Ch. 6; Searle 1983, Ch. 9). But surely we can in the same (descriptivist) way adopt the whole Kripkean picture of the referential role of names. The intentional content supposed to pick out the referent of a particular utterance of the name 'Jan Novák' (by a speaker not equipped with any independent, non-parasitic identifying knowledge) can be specified e.g. by means of the description: "the individual baptized by the name 'Jan Novák' at the beginning of the chain to which this utterance belongs".[2] The reason why the description should include not only reference to the name but also to its particular utterance, is that it serves to identify the relevant chain among (supposedly) thousands of chains in which the name 'Jan Novák' appears.

To recall Brian Loar's aphorism (from Loar 1980) , this is a way in which the causal theory of names acquires "self-consciousness". The functioning of the mechanism which links names with their referents is presented as essentially including deliberate reliance of language users on that very mechanism. This move does not weaken or blur the *external* reality of the social mechanism in question: however, from this new point of view, the speaker *himself* introduces the mechanism into the articulation of the content of his communicative act (and of the thought expressed in it). In other words, the mechanism does not apply from the outside on the speaker's act; rather, the act includes an *exploitation* of the mechanism in its articulation.

4   The social conceptual repertoire

The same kind of move, with the same motivation, should be applied to the externalist account of the role of social conceptual repertoire in the articulation of our thought contents. Let me focus on Tylor Burge's famous arthritis example, which is not discussed in Searle (1983) , unlike Burge's early externalist theory of *de re* beliefs (beliefs characterized by Burge as including ineliminable indexical elements; cf. Burge 1977), Putnam's *Twin-Earth* thought experiment and Kripke's causal theory of proper names. Nevertheless it should be, as I have just suggested, approached in the same way.

_____

[2]   Kripke himself admits such a possibility (and attempts to diminish its relevance) in a brief footnote in Kripke (1972, 88, cf. also 162).

In fact, we can, for our purposes, get along with the situation described in the first part of Burge's notoriously known example (from Burge 1979). There we learn that somebody called Bert has not fully mastered the standard notion of arthritis as it is used in our medical science. Our evidence for this is that he seriously and sincerely utters sentences like "I have arthritis in my thigh". According to Burge, this fact in itself should not prevent us from using the term "arthritis" quite straightforwardly in attributing beliefs and other attitudes to Bert. In particular, we have full right to apply disquotation on the utterance mentioned, with the result: "Bert believes that he has arthritis in his thigh." When doing so, we (provided that we use that sentence in the standard way) project the conventional meaning of the term "arthritis", i.e. the medical concept of arthritis, into the content of Bert's belief - although we know precisely about this concept that it has not been mastered by Bert.

This is quite a radical move, since in sentences like "Bert believes that he has arthritis in his thigh" the term "arthritis" occurs in the so called *oblique* position, which means that its replacement by another term identifying the same disease under another mode of presentation can change the truth value of the whole attribution. That shows that in such a sentence we are identifying the content of Bert's belief not just by relating it to certain disease but by specifying the mode of presentation under which he is thinking about it.[3] And that means that the communal conceptual repertoire fixed by the conventions of our sociolect is allowed to intervene into the apparently most internal sphere of thought content, to erase there what is incompatible with the standard and to supply what is missing. The principle seems to be that even those aspects of the concepts fixed in a given community, which have not been mastered by the subject, can participate in the articulation of his attitudes - in virtue of his having the status of a competent member of that community.

When arguing for this position, Burge (1979, 101, 114) points to an important empirical phenomenon, made popular by Putnam under the title "division of linguistic labour" (cf. e.g. Putnam 1975): individuals typically do respect the authority of experts, concerning meanings of

_____

[3] Cf. Burge (1979, 114): "Clearly oblique occurences in mentalistic discourse have something to do with characterzing a person's epistemic perspective – how things seem to him or how they are represented to him."

words and conceptual components of the contents of their attitudes. In these issues, like in many others, individuals feel obliged to the social standards even if they have not fully mastered them. To put it positively, even those concepts from communal repertoire which we have not mastered can be engaged in the articulation of our attitudes, and hence in the determination of their satisfaction conditions. This is the profit we gain from the division of linguistic and intellectual labour, if we participate in it and rely on it.

The question is whether the submissivity to experts or to social standards provides us with any argument in favour of social externalism.[4] As I have pointed out, John Searle does not discuss Burge's arthritis example in Searle (1983), even though he does comment there on the idea of the division of linguistic labour. Yet the Searlean manner of facing externalist arguments is well applicable here. If we admit that Burge's Bert, in the articulation of his belief, relies on (or defers to) the medical definition of arthritis unknown to him, the internalist should insist that this reliance should be explicitly reflected in our specification of the content of Bert's belief. The simplest and most natural way of doing so is to say: "Bert believes that he has in his thigh the disease referred to by experts in his community as 'arthritis'." Or: "Bert believes that he has in his thigh a disease satisfying the medical notion of arthritis." Or you can opt for some other formulation with the same effect – namely of directly involving Bert's deference to experts into the construal of the content of his belief.

5 Some problems

Obviously, the value of such constructions largely depends on your ability to explain their philosophical implications, to demonstrate their applicability on cases which may be found controversial, and to defend them against objections of all possible kinds. It is here that the detailed analytical work starts. Let me just mention some important issues without going into them:[5]

---

[4] An opposite view has been expressed e.g. in Searle (1983, 201f), Bilgrami (1992, 67f), Davidson (1994, 5).

[5] I have attempted to do so in Koťátko (2006a) and in more detail in Koťátko (2006b, Chap. C.III.2).

a) The analysis must be demonstrably compatible with the assumption that the deferring layperson to whom it is meant to apply can share beliefs about arthritis with an expert, despite the fact that the content of the latter's belief does not include any deference to the expert use of the term "arthritis" or to the expert notion of arthritis.

b) The same concerns the assumption that English and Czech laypersons can share beliefs about arthritis, even though the metalinguistic part of their belief contents includes, according to the analysis, reference to two different words: "arthritis" and "artritida".[6]

c) It should be made clear that the metalinguistic analysis does not turn beliefs about non-linguistic entities into beliefs about language (as T. Burge objects in 1979, 97).

d) The analysis must not imply that the sentences "I have arthritis in my thigh" and "I have in my thigh the disease referred to by experts in this community as 'arthritis'" have the same meaning.

e) The metalinguistic analysis must not lead to the consequence that statements like "Arthritis is called 'arthritis' by experts in this community" come out as necessarily true: rather, the analysis should allow us to treat them as examples of the *contingent a priori*.

f) It should be made clear that the analysis does not generate any destructive kind of regress; in particular, that it is not the case that once it is applied at a certain level its completeness requires that it also be applied at a higher level, i.e. to the expressions employed in it, etc. *ad infinitum*.[7]

g) The analysis should account for all possible distributions of the roles of the layperson and of the expert between the interpreted person and the interpreter (and of their attributions of these roles to one another). For different combinations, the structure of the belief attribution (and within it the construal of the attributed belief) will also differ.[8]

---

[6] T. Burge finds this problem quite serious (cf. Burge 1979, 96), G. Segal points to it in one of his arguments against the metalinguistic construal of belief contents (see Segal 2005, 115), H. Putnam has done the same in the discussion at the Karlovy Vary Symposium on *Swimming in XYZ* in 1998.

[7] Gabriel Segal has argued that it is the case in Segal (2000, 115). I have offered arguments against this view in Koťátko (2006, 63f.).

[8] I have attempted to say more about this in Koťátko (2010, Chap. 3).

h) The analysis should take into account the possibility of "switches" in which the interpreted person moves from one social environment to another, belonging to the same linguistic community but varying in the conceptual repertoire and in the standards governing the use of some terms.[9]

6 Some implications

The defence and elaboration of the internalist construction of the thought content with respect to these and similar issues is perhaps the most attractive part of the project – but here we have to remain on the level of the basic philosophical motivation and implications of the internalist position. In general, the kind of the construal of the thought content defended in Chapter 4 can be viewed as an internalist way of appreciating the social parameters of thought to which philosophers like Burge or Putnam have drawn our attention. If we admit that linguistic conventions or the expert usage participate (at least in some cases) in the determination of the content of our beliefs and other attitudes as radically as Burge insists, we should have no problem with the idea that in such cases the belief content itself includes an appeal to these conventions or to the expert usage. The function of this appeal is to bridge the gap between the expert meanings or communal conceptual repertoire and the actual linguistic or intellectual equipment of the individual in question. In other words, it enables to the individual subject to exploit (in the articulation of his beliefs) even those components or aspects of the standard communal conceptual repertoire which he has not mastered. To reject this while accepting the socialist claims about the nature of our language and thought (in Burge's sense) amounts to insisting that individuals are rather *objects* than *subjects* of the division of linguistic and intellectual labour; or that the integration of individuals into the community on this level is rather *external unification* than *active participation*; or that what is going on here is some kind of *self-projection* of communal concepts into the content of the subject's beliefs rather than the subject's *exploitation* of the concepts from the communal repertoire.

---

[9]   On the phenomenon of switches and its implications cf. e.g. Warfield (1987), Ludlow (1985), Ludlow (1987).

From the internalist point of view, in the Bert-like cases we have just two options. Either we feel justified to ascribe deference to social standards to the subject in question, and then we should include it in the construction of his beliefs; or we see no reasons to assume deference, and then we are obliged to do our best to identify the subject's notions, including their idiosyncratic features, on the basis of his linguistic as well as non-linguistic behaviour. The third option, namely to adopt the position of managers (or wardens) of the sociolect and of the communal conceptual repertoire fixed in it and to approach the subject's thought contents as liable to our tutelage, is simply not open to us.

The externalism-internalism dispute concerning the social nature of thought provides a good opportunity to ask again the old question what it is for an individual to *have* a concept or to *have* a belief with a certain content. From the point of view of social externalists, having a concept does not necessarily mean knowing the criteria determining its extension (and knowing the application rules of the corresponding term). That implies that an individual can have a belief with quite determinate content without being able to specify its truth-conditions: it is enough that these conditions are determined on the level of the communal conceptual repertoire. From the viewpoint of the socially sensitive internalist it is not so: we can keep the intuitive assumption that to have a concept is to know the criteria of its application (in sense of a complete set of necessary and jointly sufficient conditions), but these criteria can be *parasitic* or *deferential*. If the layperson's notion of arthritis has the form "the disease referred to by experts as 'arthritis'", it obviously gives a criterion unambiguously determining the extension, provided that the appeal to experts is satisfied, i.e. provided that in the given community there is precisely one expert meaning of the term 'arthritis' which unambiguously determines its extension. Granted this, the layperson's belief which would be properly expressed by the sentence "Arthritis is infectious" has unambiguous and internally fixed truth conditions.

From this perspective, having a belief and having a concept seems to be a more respectable role than it appears to be from the point of view of social externalists. But not necessarily much more. The deference to experts or other kinds of authorities gives them in any case some power over the content of our beliefs. For instance, a man reading in his favourite newspaper the sentence "Extremists (internalists, individualists etc.) should be made harmless" may find it appealing and

adopt a belief construed as follows: "Those who are classified by competent authorities as 'extremists' should be treated in a way referred to by competent authorities as 'making somebody harmless'." In such cases the thinking subject "retreats" from the content of his own attitude or communicative act and delegates its articulation on concepts which are in possession of others. Even this delegation can be, if we are charitable enough, interpreted as a way of the subject's involvement in (or even his control over) the articulation of the content of what he thinks or communicates – but this can hardly impress an individualist in the theory of content. The most natural thing to say about the individual from our last example is that he allows his own subjectivity to dissolve in the external collective mind (controlled by the authorities). In an extreme case, the whole conceptual work of such an individual can consist in the identification of (and appeal to) external resources of the articulation of his own thoughts and communicative acts – and then we have full right to say that the thoughts of such an individual "pass him by".

According to the individualist position I am defending here, the interpretation should treat the interpreted subject as finding himself in such an embarrassing position only if there is no other option compatible with the available evidence. In other words, the interpretation should be allowed to weaken the subject's control over the content of his own thoughts and communicative acts only under irresistible pressure of evidence. Or to put it positively: the idea of a subject maximally involved in the articulation of his attitudes and communicative acts should play the same regulative role as the idea of a coherently and rationally thinking and acting individual. To say so is to commit oneself to a certain version of the well-known *principle of charity*. I will call it the *principle of the maximization of subjectivity*, and (inspired by Kant's presentation of his categorical imperative) submit it in three equivalent formulations:

a) Do not ascribe to the interpreted person P contents (of thoughts or communicative acts) such that their articulation depends on conceptual resources beyond the capacities of P, whenever the available evidence (the communicative as well as the non-communicative behaviour of P) admits an alternative interpretation.

b) Among the alternative ways – all equally supported by the available evidence – of identifying the contents thought or communicated by the person P, choose the one which maximizes P's subjective involvement in his attitudes or acts.

c) The interpreted person should not, as a result of the interpretation, come out as more dependent on the division of linguistic and intellectual labour than is needed for making sense of his overall behaviour.

Institute of Philosophy
The Academy of Sciences of the Czech Republic, v. v. i.
Jilská 1
110 00 Praha 1
Czech Republic
kotatko@flu.cas.cz

References

Bilgrami, A. (1992): *Belief and Meaning*. Oxford: Blackwell.

Burge, T. (1977): Belief De Re. *Journal of Philosophy* 74, pp.

Burge, T. (1979): Individualism and the Mental. In: French, P. – Uehling, T. – Wettstein, H. (eds.): *Midwest Studies in Philosophy, IV: Studies in Metaphysics*. Minneapolis: University of Minnesota Press.

Davidson, D. (1994): The Social Aspect of Language. In: McGuiness, B. – Oliveri, G. (eds.): *The Philosophy of Michael Dummett*. Dordrecht – Boston – London: Kluwer Academic Publishers.

Koťátko, P. (2006): Having a Concept. In: Marvan, T. (ed.): *What Determines Content? The Internalism/Externalism Dispute*. Newcastle: Cambridge Scholar Press, 55-68.

Koťátko, P. (2010): Realismus a jeho implikace. [Realism and Its Implications] In: Zouhar, M. (ed.): *Realismus, internalizmus, individualismus. Ku knihe Petera Koťátka Interpretace a subjektivita*. Bratislava: Infopress, 11-30.

Kripke, S. (1972): *Naming and Necessity*. Oxford: Blackwell

Loar, B. (1980): Names and Descriptions. A Reply to Michael Devitt. *Philosophical Studies*, 38, 85-89.

Ludlow, P. (1985): Externalism, Self-Knowledge and the Prevalence of Slow Switching. *Analysis* 55, pp.

Ludlow, P. (1987): On the Relevance of Slow Switching. *Analysis* 57, pp.

Putnam, H. (1975): The Meaning of „Meaning". In: Putnam, H.: *Mind, Language and Reality*. Cambridge: Cambridge University Press.

Putnam, H. (1981): *Reason, Truth and History. Cambridge*: Cambridge University Press.

Searle, J. R. (1958): Proper Names. *Mind* 67, pp.

Searle, J. R. (1969): *Speech Acts. An Essay in the Philosophy of Language*. Cambridge: Cambridge University Press.

Searle, J. R. (1983): *Intentionality. En Essay in the Philosophy of Mind*. Cambridge: Cambridge University Press.

SEARLE, J. R. (1995): *The Construction of Social Reality*. London: Penguin.
SEGAL, G. (2000): *A Slim Book about Narrow Content*. Cambridge (Mass.): The MIT Press.
STRAWSON, P. F. (1959): *Individuals. An Essay in Descriptive Metaphysics*. London: Routledge.
WARFIELD, A.T. (1987): Externalism, Priviledged Self-Knowledge and the Irrelevance of Slow Switching. *Analysis* 57, pp.