MUTUAL BELIEFS AND COMMUNICATIVE SUCCESS†

Petr KOTATKO*

* Philosophical Institute of the Czech Academy of Sciences, Jilská 1, 110 00 Prague 1, Czech Republic. E-mail: Kotatko@mbox.cesnet.cz

BIBLID [0495-4548 (2000) 15: 39; p. 421-433]

ABSTRACT: The paper explores the notion of communicative success as a match between the speaker's communicative intention and the audience's interpretation. The first part argues that it cannot be generalized to all kinds of communication. The second part characterizes various types of relations between the speaker's and the audience's beliefs on which this kind of communicative success can be based. It shows that the requirements concerning agreement between these beliefs are rather modest.

Keywords: communicative intention, interpretation, utterance meaning, mutual beliefs.

CONTENTS

- 1. The notion of communicative success
- 2. The kinds of match Bibliography

1. The notion of communicative success

Here are five possible parametres of communicative success which, according 10 my view, should be kept separate, since they need not (though they typically do) coincide:

(1) the fulfilment of the speaker's intention (I_1) to perform a particular speech act (identified by propositional content and illocutionary force);

(2) the fulfilment of the speaker's intention (I₂) to be understood by the audience as having performed a particular speech act;

(3) the fulfilment of the audience's intention (I₃) to grasp the speech act

performed by the speaker's utterance;
(4) the fulfilment of the audience's intention (I₄) to grasp the speaker's

intention I_1 ; (5) the fulfilment of the audience's intention (I_5) to grasp the speaker's

intention I_2 .

THFORIA - Segunda Época Vol. 15/3, 2000, 421-433

For Grice, Searle, Davidson1 and many other authors the speaker's intentions I1 and I2 coincide, whereas I take it as a characteristic feature of human communication that they may diverge. First, one certainly can intend to perform a particular speech act without caring about being or not being interpreted by his audience as having done so: imagine a commandant saying "Fire" in the circumstances where this utterance, according to rules governing communication in the army, counts as an order to shoot. The commandant's only intention may be to fulfil his task in a certain chain of actions (understanding this task as doing whatever constitutes an order to shoot). Or he might even count on a discrepancy between the meaning of his utterance and the audience's interpretation: for example, he can intend the soldier (who is in such a position that the order unambiguously applies to him) to interpret his utterance as a warning that something is burning or an order to make fire, etc. simply because he welcomes any opportunity to punish soldiers for not obeying his orders. Such examples show that the consistency of the intention of the type I1 is not threatened if there is no intention of the type I2 or if the speech acts specified in both intentions differ in propositional content (and obviously, the same holds about force). The reason is that the speaker who has the intention I1 may count on certain meaning determining principles which do not include any reference to the audience's interpretation. In other words, the speaker's having the intention I₁ does not in itself imply that the speaker intends the audience's intention I₃ to be fulfilled, and the same for I₄ and I₅. And in the case of an audience who counts with the possibility of discrepancy between I1 and I2 or even suspects that they actually diverge, the intentions I4 and I5 in no sense coincide. It should be also clear from our example that the audience can very well intend to identify properly the speech act performed in the utterance, i.e. have the intention I_3 , without caring about speaker's intentions I_1 and I_2 . The soldier may care simply about what the rules which are in force in the given type of communication assign as a meaning to the commandant's utterance; he can even assume that the commandant is so drunk that when uttering "Fire" he does not have any intentions of the type I1 (not to speak about I2) or that his intention I1 radically diverges from the meaning of his uterance (as determined by relevant rules). In neither of these cases is there any reason to reject the ascription of a proper interpretative intention to the audience.

It is easy to continue in construing examples of this kind, not necessarily limited to as ritualized or formalized types of communication as giving military orders certainly is.² They show that one cannot do with the sim-

ple notion of communicative success, according to which the speaker and the audience succeed precisely in the same case: namely when the speaker is understood as he wanted. Since this implies that the speaker's intentions I_1 and I_2 coincide, and the same concerns the audience's intentions I_3 , I_4 and I_5 . And such a coincidence can obtain only if, first, utterance meaning is determined simply by the match between the speaker's communicative intention and the audience's interpretation; second, the speaker and the audience take the utterance meaning to be determined in such a way.

According to my opinion, this principle of the determination of utterance meaning drastically simplifies the picture of human communication. I believe that we have to admit various kinds of principles corresponding to various types of communication and differing in the kind of factors involved in the determination of utterance meanings (like linguistic conventions or social standards of justified intention ascriptions). And if we look for a general notion of utterance meaning, it should be one which allows for such a variety rather than one which imposes on us one particular principle. The Davidsonian notion of utterance meaning as constituted by the match between the speaker's communicative intention and the audience's interpretation (as presented e.g. in Davidson 1994, pp. 11-12) certainly does the latter, since the only principle of the determination of meanings of particular utterances compatible with this notion says: the utterance has the content p and force f precisely if it was so meant by the speaker and interpreted by the audience.

Now, my reluctance to the match-account of utterance meaning is not limited to the objection that it oversimplifies the picture of communication. Let us consider the following example. The speaker utters the sentence "It was raining" with the intention to assert that (in certain time in the past) it was snowing. Imagine that he uses the word "raining" for this purpose either because he is absend-minded, or he believes that it means snowing in English or in the audience's idiolect. Suppose that there is no publicly accessible evidence which would allow us to justify the claim that some of these cases obtain and hence to properly identify the speaker's intention. Nevertheless, imagine that the audience understands the utterance as an assertion that it was snowing, since he is in that very moment dreaming of a white Christmas. Then, according to the match-account of utterance meaning, the utterance meant that it was snowing, although this meaning is not identifiable by means of any public procedures (since, ex hypothesi, the speaker's intention is byond their reach). This certainly conflicts with the thesis about the public nature of meaning, which I want to keep. A possible

counter-move is to reject one of the presumptions of our example and to identify the speaker's intentions with those which are ascribable to him according to public procedures, based on public evidence. This implies a very non-traditional notion of a person's *having* intentions, beliefs, etc. for which we would need quite a strong independent motivation. But any less radical account of the public nature of propositional attitudes, allowing space for a distance between the actual intentions and public procedures for their identification, leads, when combined with the match-notion of utterance meaning, to a direct conflict with the public nature of meaning. I take this as one more reason to be suspicious about that notion.

But this certainly does not disqualify the match between the speaker's communicative intention and the audience's interpretation of the utterance as a possible parameter of communicative success: since the speaker and the audience certainly can (though, as I have argued, they need not) aim at such a match and even regard it as the only relevant criterion of success of their joint enterprise. Now, having warned against an unjustified generalization of this type of communicative situation, I wish to concentrate on it and say a bit more about its structure.

2. The kinds of match

It is essential for the type of communicative situations we will be thinking about, that the speaker's and the audience's communicative attitudes satisfy the following conditions:

(a) The speaker does not intend to perform particular speech act unless he wants to be interpreted by the audience as performing that act.

(b) The audience does not interpret the utterance as a performance of particular speech act unless he believes that it is the act the speaker wanted to perform.

This is included in the presumption that both the speaker and the audience aim at the match between the communicative intention and interpretation. Clearly, the desired match can be reached quite spontaneously, on the basis of sharing certain routine communicative practices (not to speak about cases when it is reached simply by a good luck). But the speaker can also choose words quite deliberately, on the basis of certain beliefs about how they will be interpreted by his audience on the given occasion, and the audience can interpret the utterance on the basis of certain beliefs about how the speaker wants and expects to be interpreted by him on that occasion.³ I would like to say a bit more about the possibilities which obtain here, in-

cluding those which are not typically exploited in our actual communicative practices.

2.a

First, let us consider a communication in which all participants assume that words have standard meanings in a given community. It does not matter what *they* mean by this: they may have a very vague idea which they are unable to articulate or a sophisticated theory based on, let us say, Lewisian notion of linguistic convention, or some phantastic theory based on the myth of social contract, or something even more bizzare. Important is that the notion of standard meaning plays a relevant role in their use of words and in their interpretation. But it can play this role in various ways which are worth considering:

(1) Let us imagine a communication in which the authority of the institution of standard meaning is so high that everybody uses words and interprets their utterances in a way which *he takes as* conforming to standard meanings.

(2) Or let us imagine a communication in which everybody wants to conform to the partner's idea of the standard meanings, and hence everybody uses words and interprets their utterances on the basis of his belief about what the partner takes to be the standard meanings of words.

(3) Or let us imagine a communication in which every interpreter wants to conform to the speaker's beliefs about what the interpreter takes the standard meanings of words to be: and similarly for the speaker's use of words.

Obviously, we can continue in this way, proceeding to more and more complex forms of communicative *mutuality*. In each of these cases, mutuality is limited by the fixed order of beliefs which are relevant for the choice of words and for their interpretation. It should be clear that this limitation is incompatible with the constraints (a) and (b) imposed by the match-account of communicative success on the communicative attitudes. Let us say that the relevant beliefs are of the order n. In that case the audience can still have also an n+1 order belief about the speaker's n-order belief determining the speaker's choice of words. But then the interpretation should be based on this n+1 order belief, otherwise the audience would not interpret the utterance in the way which he takes to have been intended and expected by the speaker. And similarly for the speaker's choice of words. The only type of communication which is compatible with the match-

account of communicative success is one which does not impose any limits on the communicative mutuality.

In the rest of the paper I want to say something about communication of this type: my aim is to characterize the kinds of beliefs which can be decisive for the choice of words and for their interpretation and to specify the kind of match between them which must obtain in order for there to be communicative success of the type we are interested in, i.e. success in the sense of a match between communicative intention and interpretation.

Let us first distinguish three kinds of match between S's and A's beliefs. The first one, which I propose to call *full agreement* in beliefs, obtains if all the S's beliefs representing A's relevant beliefs (beliefs of the kind to be specified bellow) are true and *vice versa*. For example:

 $B_Ap \& B_AB_Sq \& B_AB_SB_Ap$ etc. $B_Sq \& B_SB_Ap \& B_SB_AB_Sq$ etc.

[p is the proposition that the sentence X conventionally means M_1 , q is the proposition that the sentence X conventionally means M_2 ; B_Ap reads: A believes that p; B_AB_Sq reads: A believes that S beli

Notice that once we ascribe to A the belief $B_A p$, we must have p after all occurences of B_A in both series, if full agreement is to obtain: since the full agreement requires that S has the belief $B_S B_A p$ (if he has any belief of the second order at all), and this is truly represented in A's belief $B_A B_S B_A p$ (if A has any belief of the third order), which is truly represented in S's belief $B_S B_A B_S B_A p$, etc. And analogically with $B_S q$. So: once we have fixed the first two conjuncts in one series, or any neighbour conjuncts in any of the two series, this determines the only possible development of both series.

Of course, it is not a priori given where these series should stop (i.e. what the order of the highest order beliefs should be). The full agreement does not even require that both series stop at the same level, but the difference between the order of S's and A's highest order beliefs can be at most 1: otherwise one participant ascribes to the other a belief which he does not have. In principle, any belief (independently of the order) can be missing. But in case of full agreement, the gaps must be in a certain correlation: if there is a gap in S's series at the level n (i.e. it does not include any n-order belief), the A's series must have a gap at the level n+1, since otherwise one participant ascribes to the other a belief of the order he does not have. But then S's series must have a gap at the level n+2, and so we get a zig-zag correlation between the locations of the gaps. Let us say that A does not

have any first-order belief: then the gaps (marked by italics) must be distributed in the following way:

 $B_A p \& B_A B_S q \& B_A B_S B_A p \& B_A B_S B_A B_S q$ etc. $B_S q \& B_S B_A p \& B_S B_A B_S q \& B_S B_A B_S B_A p$ etc.

If there is only a one-sided agreement of the kind mentioned, for example, if all S's beliefs representing A's relevant beliefs are true, but not vice versa, I propose to speak about a partial agreement between S's and A's beliefs.

The strongest kind of agreement I can think of can be called *exhaustive* agreement between Ss and A's beliefs. It obtains if and only if not only all Ss beliefs about A's relevant beliefs are true and *vice versa*, but S has true beliefs about all A's relevant beliefs (i.e. there is no relevant belief of A which would not be truly represented in Ss belief), and *vice versa*. The exhaustive agreement necessarily includes an infinite series of beliefs on both sides, as the following consideration shows:

Let us suppose that S believes that p. The exhaustive agreement requires that A knows this, which means that A must believe that S believes that p. If this is fulfilled, then it is something which must be known by S (otherwise the exhaustive agreement would not obtain), which means that S must believe that S believes that S belief th

 $B_sp \& B_Aq \& B_sB_Aq \& B_AB_sp \& B_sB_AB_sp \& B_AB_sB_Aq \& ... etc.$ ad infinitum.

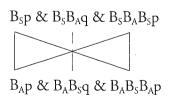
Like in the case of full agreement, once we have the first two conjuncts, the rest follows. If the first order S's belief is B_Sp , then in all the conjuncts (no matter whether they represent A's or S's beliefs), which have B_S in the end position, the last embedded proposition must be p. But unlike in the case of full agreement, there cannot be any gaps here and the series is infinitely expanded.

Our notion of exhaustive α reement in beliefs is reminiscent of the notion of mutual belief or knowledge as defined in Schiffer (1972) (and analogical notion of common knowledge as defined in Lewis 1969), but there is a difference in that S and A need not share any belief. In our particular example, we do not have the case of mutual belief that p, neither that q: we have different propositional contents of the beliefs ascribed to S and A at the first order level and there is no pair of beliefs of S and A such

that they are of the same order and share the last embedded proposition. Let us speak in such a case (when we have an exhaustive agreement in beliefs but not a mutual belief in Schiffer's sense) about correlated belief. No doubt, the correlated belief has an essential common feature with Schiffer's mutual belief: both are infinitely expanded series of beliefs of two participants of growing order, such that any belief of n+1 order of one participant is a belief about an n order belief of the other participant (where $n \ge 1$). Hence they are characterized by certain mutuality (mutual relatedness of the belief contents) and this can provide a basis for certain coordinated action, e.g. communicative enterprize directed on both side to the match between communicative intention and interpretation.

I have made these distinctions just in order to say that neither match of the kinds mentioned is necessary for there being communicative success in our sense, i.e. the match between communicative intention and interpretation. What is necessary is just that the belief on which the speaker bases his choice of words and the belief on which the audience bases his interpretation have the same last embedded proposition. Let us call the beliefs on which either the choice of words or the interpretation are based, effective beliefs. They need not be identical with the highest order beliefs. For instance, S's highest order belief may be $B_sB_AB_sq$, but S may believe that the belief he ascribes there to A (i.e. B_AB_sq) will not play any role in A's interpretation of the utterance. The reason can be that S does not take A to be sophisticated enough to exploit in his interpretation a belief of such a complexity (though he is sophisticated enough to have it); then S will base his choice of words on his second-order belief B_sB_Aq . And analogically for A.

So, communicative success in our sense requires that the *effective* beliefs have the same basic propositional contents (i.e. the same last embedded proposition). It does not require that these beliefs be true. In general, the communicative enterprise can be successful (in the sense of the match-account of communicative success) even if all of S's and A's relevant beliefs about meanings and about the partner's attitudes, including the effective believes, are false. For instance, let us have the following two series of beliefs:



Here the second and third order beliefs of S and A are obviously false. Let us also suppose that p is not the case (i.e. the sentence X does not conventionally mean M_I) so that both first order beliefs are false as well. But quite independently on this, there are no less then five possibilities of communicative success: the communication is successful if any of the five pairs of beliefs connected by the lines becomes effective (i.e. determines S's choice of words and A's interpretation).

2.b

While the conclusion we have just reached may be found surprising, the following one was easy to predict at the very beginning: communication can be successful (in sense of the match-account) even if S and A do not use and interpret words in the standard way (they need not have any first order beliefs and if they have them these beliefs may be false), neither uses and interprets words on the basis of their beliefs about standard meanings (again, S and A need not have any first order beliefs and if they have them these beliefs may be ineffective), neither uses and interprets words on the basis of their beliefs about what the partner takes to be standard meanings (they need not have any second order beliefs and if they have them these beliefs may be ineffective), etc. For it is still possible that the choice of words and their interpretation are based on some higher order beliefs. What is important is that in the kind of beliefs we have been considering till now the notion of standard meaning still plays an essential role: it is not only present in their propositional content, but determines the relations between them. Hence the *elimination* of this notion will inevitably change the entire structure. The first order beliefs are now S's belief about how he will be interpreted if he utters certain words and A's belief about what S means if he uttters certain words. Plainly, these beliefs immediately determine the choice of words and their interpretation. Although there can be also higher-order beliefs, these cannot "defeat" or override the first order beliefs in the choice of the communicative strategy: they can only justify them. For example, S's belief that if he utters certain words he will be interpreted so and so can be based on S's belief that A believes that if S utters those words he means so and so. And this S's belief can be based on his belief that A ascribes to him the belief that if he utters certain words he will be interpreted so and so, etc.

Let me put this a bit more precisely. Once we eliminate the notion of the standard meaning we have to replace the propositions p and q concern-

ing what is the standard meaning of the sentence X by propositions concerning what is meant by the utterance of that sentence on a given occasion or how the utterance will be interpreted. This means that we have to take into account four propositions instead of two, even if we still consider only two candidates for utterance meaning, namely M_1 and M_2 :

 p_i : if S utters X (on a given occasion) he means M_i ;

 p_2 : if S utters X (o.g.o.) he would be interpreted by A as meaning M_j ;

 q_1 : if S utters X (o.g.o.) he means M_2 ;

 q_2 : if S utters X (o.g.o.) he would be interpreted by A as meaning M_2 .

Obviously, S can believe that his utterance of X would be interpreted by A as meaning M_I (B_Sp_2) without having the second-order belief that A believes that by uttering X, S would mean M_I ($B_SB_Ap_1$), not to speak of higher order beliefs ($B_SB_AB_Sp_2$, etc.). But S's first order belief, together with his beliefs concerning the obtainment of the conditions (a) and (b) (cf. beginning of the part 2), imposes constraints on his higher order beliefs: thus, if he takes the condition (b) as fulfilled, the only second-order belief compatible with B_Sp_2 is $B_SB_Ap_1$. And under the same presumption the latter belief justifies the former. This can be generalized in the following series based on the relation of justification (\leftarrow):

 $B_{S}p_{2} \leftarrow B_{S}B_{A}p_{1} \leftarrow B_{S}B_{A}B_{S}p_{2} \leftarrow B_{S}B_{A}B_{S}B_{A}p_{1} \leftarrow B_{S}B_{A}B_{S}B_{A}B_{S}p_{2} \text{ etc. } \textit{ad infinitum.}$

And analogically on the part of A:

 $B_A p_1 \leftarrow B_A B_S p_2 \leftarrow B_A B_S B_A p_1 \leftarrow B_A B_S B_A B_S p_2 \leftarrow B_A B_S B_A B_S B_A p_1 \text{ etc. } \textit{ad infinitum.}$

Taken together, these two series exemplify what we have called *full agreement* in beliefs (all S's beliefs representing A's beliefs are true and *vice versa*). And if we take them as infinitely expanded, they also exemplify what we have called *exhaustive agreement* in beliefs (in addition to what has been said, all S's relevant beliefs are truly represented in A's beliefs and *vice versa*).

These justification relations depend on the beliefs concerning the fulfilment of the conditions (a) and/or (b), on the beliefs concerning the partner's beliefs regarding this matter, etc., in a rather interesting way. At first sight, one may be tempted to say that the two infinite series together obtain only on the assumption that S and A mutually believe (in sense of Schiffer, 1972) that the conditions (a) and (b) are fulfilled: and that for any particular case of justification some fragment of this infinitely expanded belief is needed. But this is not the case: the structure we are confronted with here is different from Schiffer's mutual belief (or its fragments). To see this, let us suppose that S who wants to make an utterance with the meaning M_I decides to utter X on the basis of the belief that A will interpret his utterance (on a given occasion) as meaning M_I : in that case the condition (a) is fulfilled. This belief can (though it need not) be based on Ss belief that A believes that if S utters X, he means M_I : this requires that S takes the condition (b) as fulfilled. And this latter belief can (though it need not) be based on Ss belief that A ascribes to S the belief that if he utters S, he will be interpreted as meaning S0 this requires that S1 takes S2 takes S3 takes S4 as taking the condition (a) as fulfilled. Etc.

This can be schematically represented as follows:

[Here the double arrow \$\extstyle{\textstyle{\textstyle{1}}}\$ represents the relation of being based on: it points from a justification relation to a belief on which the justification in question is based. The symbol \$\textstyle{\textstyle{1}}\$ represents a rather complex relation: if the acting subject opts for some action (in our particular case makes some utterance) on the basis of the belief ascribed to him in the formula at the bottom, then the state of affairs specified at the top obtains.]

And analogically on the part of A. A can interpret Ss utterance of X as meaning M_I because A believes that S intends to be so interpreted: this means that the condition (b) is fulfilled. This belief of A can (though need not) be based on his belief that S uttered X believing that A would interpret him as meaning M_I : this requires that A takes the condition (a) as fulfilled. And this latter belief of A can be based on his belief that S takes him as believing that S, when uttering S, means S0 this requires that S1 takes S2 as taking the condition (b) as fulfilled. Etc. Let us represent this as follows:

So, if we put together S's and A's beliefs required for there being the infinite justification series specified above, we get an instance of *exhaustive* agreement in beliefs, which is at the same time not an instance of Schiffer's mutual belief: in 2.a we have decided to call this *correlated belief*.

 $B_sb \& B_Aa \& B_SB_Aa \& B_AB_Sb \& B_SB_AB_Sb \& B_AB_SB_Aa \& B_SB_AB_SB_Aa \& \& B_AB_SB_AB_Sb$ etc. ad infinitum.

To conclude: the main distinction between the two types of communication we have been considering in sect. 2.a and 2.b is that in the latter case (i.e. in communication where the beliefs relevant for the choice of communicative strategy do not include the notion of standard meaning) the first order beliefs play a privileged role. The match which constitutes communicative success obtains simply if S and A have corresponding beliefs of the first order. In the choice of words and in their interpretation these beliefs cannot be overridden by any higher-order beliefs. The only way in which the higher-order beliefs can contribute to the choice of communicative strategy is to participate in the justification of the first order beliefs (granted (a) and (b) and a relevant fragment of the coordinated belief concerning (a) and (b)).

Notes

- † This paper is based on the discussion I have had at the meeting of the Aristotelian Society in April 1998: I would like to thank its participants for inspiring criticism and valuable suggestions. The content of the present paper partially consists in correction or development of the views published in Kotatko (1998).
- ¹ Cf. the construction of Grice's definitions of utterer's meaning in his classical articles collected in Grice (1989), the Gricean element in Searle's famous definition of promise (condition (8)) and of other speech act types to be found in Searle (1970) and Davidson's account of utterance meaning as presented e.g. in Davidson (1986, 1994).
- ² I have given another one in Kotatko (1998, p. 234 f).
- ³ Such cases can be described in terms of Davidson's 'prior' and 'passing theories' which the speaker has for the given audience, and the other way round. Cf. Davidson (1986).
- ⁴ But it also does not belong to the nature of these series that they should be expanded *in infinitum*, unlike in case of Schiffer's *mutual belief* (cf. Schiffer 1972) or of what will be bellow called *coordinated belief*.
- ⁵ This is how I have mistakenly put it in Kotatko (1998, p. 233).

BIBLIOGRAPHY

Davidson, D.: 1986, 'A Nice Derrangement of Epitaphs', in E. Lepore (ed.): Truth and Interpretation. Perspectives on the Theory of Donald Davidson, Oxford, Basil Blackwell.

Davidson, D.: 1994, 'The Social Aspect of Language', in B. McGuiness and G. Oliveri (eds.): The Philosophy of Michael Dummett, Dordrecht-Boston-London, Kluwer Academic Publishers.

Grice, P.: 1989, Studies in the Ways of Words, Harvard University Press.

Kotatko, P.: 1998, 'Two Notions of Utterance Meaning', *Proceedings of the Aristotelian Society* III, 225-239.

Lewis, D.: 1969, Convention. A Philosophical Study, Harvard University Press.

Schiffer, S.: 1972, Meaning, Oxford, Oxford University Press.

Searle, J.: 1970, Speech Acts: An Essay in the Philosophy of Language, Cambridge University Press.

Petr Kotatko is the Head of the Department of Analytic Philosophy at the Philosophical Institute of the Czech Academy of Sciences in Prague and a co-ordinator of the Karlovy Vary Symposia on Analytic Philosophy. His writings in the philosophy of language include: Language and the Theory of Meaning (1994), Meaning and the Third Realm (1995), Descriptions and Reference (1997), Meaning and Communication (1998).