

# Lewis, Language, Lust and Lies

Max Kölbel

Universidad Nacional Autónoma de México

David Lewis has tried to explain what it is for a possible language to be the actual language of a population in terms of his game-theoretical notion of a convention. This explanation of the actual language relation is re-evaluated in the light of some typical episodes of linguistic communication, and it is argued that speakers of a language do not generally stand in the actual language relation to that language if the actual language relation is explicated in Lewis's way. In order to avoid these counterexamples, an alternative account of the actual language relation is proposed which makes use of Lewis's notion of convention in a different way.

'Ti amo', said Romeo to Juliet, thereby defying convention, for a Montague boy just did not declare love for a Capulet Girl. There was one sort of convention, however, Romeo humbly conformed to: the linguistic conventions prevalent in Verona at the time. Had different linguistic conventions prevailed, he would have emitted a different sound, for example the sound 'I love you'.

Most people agree that language is, in some sense, conventional. David Lewis has tried to make precise in what sense. In his book *Convention*, and in the later article 'Languages and Language', he attempts to explicate the notion of convention and then to characterize linguistic conventions as conventions of a particular sort, namely *conventions of truthfulness and trust in a language*, following a suggestion by Erik Stenius (1967).

As Margaret Gilbert (1981, 1983) has shown, there may be doubts as to whether the notion Lewis defines is indeed our common concept of convention. But independently of whether Lewis has captured our concept of convention, his account does seem to provide an attractive game-theoretical explanation of linguistic behaviour.<sup>1</sup> More importantly, though, Lewis's account of linguistic conventions offers a promising explication of the so-called 'Actual Language Relation' (ALR), i.e. of the relation that holds between a merely possible language  $L$  and a population when  $L$  is the (or a) language used by that population. In this paper, I re-evaluate both the purported explanation of linguistic behaviour and the attempted explication of the ALR. It will turn out that, in failing to explain many typical cases of linguistic behaviour, Lewis's account also fails as an explication of the ALR. However, I shall also show how the Lewisian account of linguistic convention might be repaired, by viewing linguistic conventions not as Lewisian

conventions of ‘truthfulness and trust’, but as Lewisian conventions of a different sort.

## I. Lewis’s Account of Linguistic Convention

I begin by summarizing Lewis’s account. Like Lewis, I concentrate on declarative sentences and the conventions governing them. My treatment can be generalized, however, to cover also sentences with different illocutionary force, such as interrogative and imperative sentences.<sup>2</sup>

First let me briefly introduce Lewis’s account of convention. Here is an example from Lewis’s book:

In . . . Oberlin, Ohio, until recently all local telephone calls were cut off without warning after three minutes. Soon after the practice had begun, a convention grew up among Oberlin residents that when a call was cut off the original caller would call back while the called party waited. Residents usually conformed to this regularity in the expectation of conformity by the other party to the call. In this way calls were easily restored, to the advantage of all concerned. (p. 43)

This example has all the marks of a convention in Lewis’s sense. The regularity is one solution, among several, to a recurrent coordination problem among the agents concerned. For each time a call is interrupted, the two parties must coordinate their actions if they want to restore the call. If both call back or both wait, they won’t be able to restore the connection. One must wait while the other calls back. It doesn’t matter which party waits and which calls back. All that is important is that they coordinate.<sup>3</sup> But how? The described convention is the solution to this recurrent problem. As it is common knowledge among residents that the called party will wait while the original caller calls back, the original caller expects the other to wait and therefore calls back, while the called party expects the other to call back and therefore waits. Many different regularities could have served the same purpose more or less equally well. For instance the regularity that it is always the party senior in age who calls back.

Let me summarize. A convention is a regularity in the behaviour (in the widest sense) of the agents of a population. This regularity arises from the common interest of the agents to coordinate their actions and is upheld because each agent expects the others to conform and prefers to conform if the others conform. Finally, there are potential alternative regularities which could also have effected coordination. This is Lewis’s full definition:

A *regularity* *R* is a convention in a population *P*, iff, within *P*, the following six conditions hold (with at most a few exceptions to the ‘everyone’s’):

- (1) Everyone conforms to *R*.
- (2) Everyone believes that the others conform to *R*.
- (3) This belief that the others conform to *R* gives everyone a good and decisive reason to conform to *R* himself.
- (4) There is a general preference for general conformity to *R* rather than slightly-less-than-general conformity – in particular, rather than conformity by all but any one.
- (5) *R* is not the only regularity meeting the last two conditions.
- (6) The various facts listed in (1)–(5) are matters of common (or mutual) knowledge: they are known to everyone, it is known to everyone that they are known to everyone, and so on. This knowledge may be merely potential: knowledge that would be available if one bothered to think hard enough.<sup>4</sup>

Now let me give a miniature example of a *linguistic* convention on Lewis's account. Our agents are the citizens of Verona, among them Romeo and Juliet. The convention consists of regularities concerning the conditions under which certain sentences are uttered (call them 'utterance regularities') and of regularities of forming beliefs in response to utterances ('response regularities'). In order to conform to the utterance regularities one must try to utter the sentence 'ti amo' only if one loves the person addressed and to utter the sentence 'ti odio' only if one hates the person addressed. Agents who conform to the response regularities tend to come to believe that they are loved by a speaker, if the speaker, in addressing them, utters 'ti amo', and to come to believe that they are hated by such a speaker if he utters 'ti odio'.

Within certain limits, the prevalence of such a convention allows Romeo and Juliet to communicate. For example, if Romeo loves Juliet and wants her to know that, he can utter 'ti amo' (thereby addressing her). For if Juliet conforms to the convention, she'll come to believe that Romeo loves her upon hearing his utterance. And he can reasonably expect her to conform, because he knows that she prefers to conform if he does, and that she expects him to conform, because he prefers to conform if she does. Alternatively, Juliet might decide, for obvious reasons, to utter 'ti odio', thereby addressing Romeo. Known to Juliet, the effect of this will be that Romeo comes to believe that Juliet hates him, because he prefers to conform if she does, and he expects her to conform.

According to Lewis, the miniature language game I have just described is just a small part of the larger language game of speaking Italian. In order to generalize this view, let's describe the miniature convention once again. First, we define a language called Lovetalk. Lovetalk is a function that assigns a truth condition (relative to a context of utterance) to each of the two sentences 'ti amo' and 'ti odio'. According to Lovetalk, an utterance of 'ti amo' is true if and only if the utterer loves the person he addresses in his utterance. An

utterance of ‘*ti odio*’, on the other hand, is true in Lovetalk iff the utterer hates the addressee. Now Lewis’s suggestion is that a population *uses* Lovetalk, just if a convention of *truthfulness and trust* in Lovetalk prevails in that population. To be truthful in Lovetalk is to try to utter a sentence of Lovetalk only if the utterance would be true in Lovetalk. To be trusting in Lovetalk is to have a tendency to respond to the utterance of a Lovetalk sentence by coming to believe that it is true in Lovetalk.

Now the generalization: any function from a domain of sentences into a range of truth conditions counts as a (possible) language.<sup>5</sup> Natural languages, however, are much more complicated and have many more sentences than Lovetalk. Languages often contain infinitely many sentences, so that their truth conditions need to be specified recursively. For example, let us think of Italian as a function that assigns a truth condition to each (declarative) Italian sentence. Now, for a population to use Italian, i.e. to have Italian as their language, is for it to have the convention of truthfulness and trust in Italian. In other words, the fact that the Italians speak Italian consists in the fact that there prevails among them the convention to try and utter an Italian sentence only if it is true in Italian, and to come to believe the truth-in-Italian of Italian sentences uttered.

Essentially, this is Lewis’s proposal. According to his explication of the ALR, a population *P* speaks, or uses, a language *L* just if there prevails among the members of *P* a convention of truthfulness and trust in *L*. He thereby also offers a neat way in which linguistic behaviour can be explained in terms of the beliefs and desires of speakers and hearers.<sup>6</sup>

## II. Lewis’s Linguistic Conventions Clarified

Now I would like to look at conventions of truthfulness and trust in more detail. According to Lewis, to be truthful in a language *L* is to *try* never to utter any sentences of *L* that are not true in *L*. To be trusting in *L* is to *tend* to respond to another’s utterance of any sentence of *L* by coming to believe that the uttered sentence is true in *L*. But what exactly counts as conformity to a convention of truthfulness and trust? In order to conform, does one need to utter only sentences that are in fact true, or does the attempt alone suffice? Does one need actually to come to believe in the truth in *L* of a sentence uttered, or does the tendency to do so suffice?

To get a clear picture, let us distinguish two different conventions concerning *L*. First, there may be a convention of *reliability and reliance* in *L*. This is the conventional regularity of reliably making only true-in-*L* utterances of sentences in *L*, and of relying on the truth-in-*L* of utterances heard. If they conform, speakers are reliable and hearers are reliant. This convention is violated whenever a false-in-*L* utterance is made and whenever

an utterance is not believed to be true-in-*L* by a hearer. Secondly, consider a convention of *sincerity and trust* in *L*. This is the conventional regularity of making an utterance only if one *believes* it to be true-in-*L* and of responding to utterances by coming to believe that their utterer believes them to be true-in-*L*. Thus, a convention of sincerity and trust in *L* is violated whenever someone makes an utterance of a sentence of *L* that he does not himself believe to be true-in-*L*, and whenever a hearer does not respond to an utterance by coming to believe that its utterer believes it to be true-in-*L*.

In order to evaluate and illustrate these two proposals, I shall test them against some fairly typical situations of linguistic communication. First, consider Case A, a case of true love:

*Case A: True Love*

Romeo utters 'ti amo', addressing Juliet.

sincere        He believes his utterance to be true in Italian.

reliable       His utterance is true in Italian.

reliant        Juliet comes to believe that his utterance is true in Italian.

trusting       Juliet comes to believe that Romeo believes his utterance to be true in Italian.

Romeo, as he addresses Juliet, utters 'ti amo'. He utters this sentence because he believes that he loves Juliet and wants her to know that. Moreover, Juliet comes to believe, upon his utterance, that Romeo loves her, because she regards him as reliable in this matter and trusts him to be sincere. Thus, they communicate successfully and information flows from Romeo to Juliet. She comes to believe something true that he did, but she didn't believe beforehand.

In Case A, both Romeo and Juliet conform to a convention of reliability and reliance in Italian. They also conform to a convention of sincerity and trust in Italian. Both conventions can explain their linguistic behaviour. But they do so in different ways. We can explain why Romeo uttered 'ti amo' by saying that he expected Juliet to conform to reliance and therefore preferred to conform to reliability himself, or by saying that he expected Juliet to conform to trust and therefore preferred to conform to sincerity. However, their latter conformity, to sincerity and trust, would not yet explain the flow of information that took place, i.e. of the information that Romeo loves Juliet. For this information to flow, Romeo needs not only to *believe* the sentence he utters to be true, but that sentence must also *be* true. And Juliet needs not only to trust that Romeo himself believes what he says, but she must also *rely* on his judgment and therefore form the belief that he loves her. Thus, a

convention of reliability and reliance can explain the flow of this information on its own, while a convention of sincerity and trust needs to be supplemented with additional assumptions to explain this.

Now consider Case B:

*Case B: Mere Lust*

Romeo utters 'ti amo', addressing Juliet.

sincere        He believes his utterance to be true in Italian.

unreliable    His utterance is not true in Italian.

unreliant     Juliet does not come to believe his utterance to be true in Italian.

trusting      Juliet comes to believe that Romeo believes his utterance to be true in Italian.

In Case B, Romeo mistakes lust for love. He does *believe* that he loves Juliet, but his belief is false. Juliet, on the other hand, does believe that Romeo's utterance is sincere, but she refuses to rely on his judgement. She does trust that he sincerely *believes* that he loves her, but this doesn't entice her to believe that he loves her. In this case, neither party conforms to reliability and reliance, but both conform to sincerity and trust. Nevertheless, Case B is a fairly typical case of linguistic behaviour, and some sort of communication takes place. Although no information with the content that Romeo loves Juliet flows, they still coordinate and communicate in some way. If we suppose that Juliet didn't believe that Romeo believed that he loved her beforehand, then Romeo at least succeeded in conveying this belief of his. The coordination that takes place in this second case cannot be explained by a convention of reliability and reliance in Italian, but it *can* be explained by a convention of sincerity and trust in Italian. Thus, in the cases considered, the convention of sincerity and trust provides a far better explanation of linguistic communication, for it applies in a case in which a convention of reliability and reliance does not apply.<sup>7</sup>

These considerations about the two kinds of convention can be generalized. Whenever a speaker genuinely, i.e. intentionally, conforms to reliability in a language *L*, he also conforms to sincerity in *L*. For how else would he make sure to make only true-in-*L* utterances than by uttering only those he believes to be true in *L*. Unintentional utterances of true sentences the speaker does not believe to be true are possible, but cannot play a role in explaining communication as a conventional activity. For conventional coordination relies essentially on the mutual replication of reasoning by agents and the resulting mutual expectations of intentional behaviour. Similarly, whenever a hearer

genuinely conforms to reliance in  $L$ , i.e. responds to an utterance by coming to believe that it is true in  $L$ , he also conforms to trust in  $L$ . For consider a hearer who responds to the utterance of a sentence of  $L$  by coming to believe that it is true in  $L$ , but does not also come to believe that the speaker believes it to be true in  $L$ . Then he must have independent reasons for coming to believe that the sentence is true in  $L$ , and thus does not genuinely conform to reliance in  $L$ .

But while every case of genuine conformity to reliability and reliance in a language  $L$  is also a case of conformity to sincerity and trust in  $L$ , there are many cases of conformity to sincerity and trust in  $L$  that are not cases of conformity to reliability and reliance in  $L$  – e.g. Case B. Thus, conventions of sincerity and trust offer a far better because more general explanation of linguistic behaviour.

But the failure of conventions of reliability and reliance as far as the explanation of behaviour in cases like Case B is concerned also throws doubt on an explication of the ALR in terms of such conventions. For according to such an explication a population  $P$  uses a language  $L$  just if there prevails a convention of reliability and reliance in  $L$  among  $P$ . In order for such a convention to prevail, however, the members of  $P$  must quite generally conform to reliability and reliance in  $L$ , and that conformity must be a matter of common knowledge. But since cases like Case B are frequent, users of a language  $L$  do not generally conform to a regularity of reliability and reliance in  $L$ , nor is it common knowledge among them that they do so. For example, speakers of Italian neither utter ‘*ti amo*’ only if that utterance is true, nor do they believe that the others utter it only if it is true. There are too many violations of reliability and reliance among users of a language  $L$ , so that they do not count as having a convention of reliability and reliance in  $L$  – there are more than ‘a few exceptions to the ‘everyone’s’ in Lewis’s definition of convention (refer back to § I for the full version of that definition).

One might argue that whether conventions of reliability and reliance do prevail among speakers of a language is a question of how pedantically we interpret ‘a few exceptions’ in Lewis’s definition. But it is important to keep in mind that it is essential to the stability of a conventional regularity that violations are kept to a minimum. The more violations agents believe there to be, the less reason they have for conforming themselves. A convention-like regularity with too many exceptions would therefore not perpetuate itself. For that reason the restriction to ‘a few’ violations cannot be relaxed. Italians do not have a convention of reliability and reliance in Italian. There simply are too many violations: too many untrue-in-Italian utterances and too many utterances not believed by their audience to be true-in-Italian.<sup>8</sup>

In this section, I have set out to make more precise what sort of convention Lewis should be interpreted as having in mind, when he claims that users of a language  $L$  try to make only utterances that are true in  $L$ , and tend to come to believe in the truth of utterances they hear. We have seen that conventions of

sincerity and trust fare much better than conventions of reliability and reliance, because the latter falsify the explication of the ALR and fail in the explanation of many typical cases of linguistic behaviour.<sup>9</sup>

### III. Many More Conventions

Now that we know how Lewis's account is best understood, it is time to be a bit more realistic about Romeo and Juliet. After all, it is conceivable that Romeo intentionally utters 'ti amo' although he doesn't believe that he loves Juliet. In other words, Romeo might lie to Juliet in order to viciously deceive her. But Juliet might not be so naïve. She might see through him and realize that he wants to give the impression that he believes he loves her, although he very well knows that it isn't love. This is what happens in Case C:

#### *Case C: Mean Lies*

Romeo utters 'ti amo', addressing Juliet.

- |            |   |
|------------|---|
| cagey      | He wants to give the impression that he believes his utterance to be true in Italian.                                 |
| insincere  | He does not believe his utterance to be true in Italian.  |
| unreliable | His utterance is false in Italian.  |
| unreliant  | Juliet does not come to believe his utterance to be true in Italian.  |
| no trust   | Juliet does not come to believe that Romeo believes his utterance to be true in Italian.                              |
| cagey      | Juliet comes to believe that Romeo wants to give the impression that he believes his utterance to be true in Italian. |

In this case, neither Romeo nor Juliet conform to a convention of sincerity and trust in Italian. Nevertheless, they are engaging in a fairly typical form of linguistic behaviour. They use words with their ordinary meaning (one wouldn't normally say that they are violating any rules of language). They also succeed in communicating something by coordinating their actions.<sup>10</sup> For Romeo hopes that Juliet will recognize his desire to give the impression that he believes that he loves her. And Juliet satisfies this desire, for she recognizes that he wants to give this impression. Of course, Juliet realizes that the impression he hopes to give would be a false impression and therefore she doesn't reward Romeo's further hope of deceiving her.



Perhaps there is yet another kind of convention that can explain this form of linguistic communication. Consider a convention of *cageyness* in Italian. This is the conventional regularity of uttering an Italian sentence only if one wants to give the impression that one believes the utterance to be true in Italian; and of responding to utterances of Italian sentences by coming to believe that its utterer wants to give the impression that he believes it to be true in Italian. Even when Romeo lies to Juliet, and even when Juliet doesn't trust Romeo to be sincere, they still conform to this regularity of cageyness.

If we again compare conventions of cageyness in a language  $L$  generally with conventions of sincerity and trust in  $L$ , we come to a similar result as in the comparison of sincerity and trust with reliability and reliance. Generally, when we can explain a piece of linguistic behaviour by a convention of sincerity and trust in  $L$ , we can also explain it by a convention of cageyness in  $L$ , and any case of (intentional) conformity to sincerity and trust in  $L$  is also a case of intentional conformity to cageyness in  $L$ . For suppose someone utters a sentence  $s$  of  $L$  because he wants to conform to sincerity and trust in  $L$ , i.e. because he expects his audience to respond by coming to believe that he, the speaker believes  $s$  to be true in  $L$ . Then this speaker also wants to give the impression that he believes  $s$  to be true in  $L$ . And suppose a hearer conforms to sincerity and trust in  $L$  because he expects an utterer to make only utterances he believes to be true in  $L$ . Then this hearer also conforms to cageyness in  $L$ , for he also believes that this utterer wants to give the impression of believing his utterances to be true. On the other hand, there are many cases of lying, cases of lacking trust, and of both combined where only a regularity of cageyness, but not a regularity of sincerity and trust, is being conformed to and can therefore explain the sort of linguistic communication that takes place. Thus, again, conventions of cageyness seem better candidates for our conventions of language than conventions of sincerity and trust.

By now, we ought to be suspicious. What prevents the succession leading from reliability and reliance via sincerity and trust to cageyness from going on indefinitely? We can think of conventions of uttering sentences only if one wants to give the impression that one wants to give the impression that one believes it to be true in some language. And of responding to utterances by coming to believe that their utterer wants to give the impression that he wants to give the impression that he believes the uttered sentence to be true in some language. We just need to insert another 'wants to give the impression that he' and thereby get a new type of convention (see Figure 1).

Let us call the resulting series of conventions that combine speaker and hearer regularities with same index numbers '<sh<sub>1</sub>, sh<sub>2</sub>, . . . , sh<sub>n</sub>>' (see Figure).<sup>11</sup> Until now, it always turned out that an sh<sub>*i*</sub> could explain more cases than an sh<sub>*i*-1</sub>, and an sh<sub>*i*</sub> was violated less frequently than an sh<sub>*i*-1</sub>. On the other hand, in those cases that could be explained both by an sh<sub>*i*</sub> and an sh<sub>*i*-1</sub>,

Speaker regularities In order to conform, a speaker must:	Hearer regularities In order to conform, a hearer must:
<p><math>s_1</math> (reliability): make only true utterances</p> <p><math>s_2</math> (sincerity): make only utterances he believes to be true</p> <p><math>s_3</math> (cageyness): make an utterance only if he wants to give the impression that he believes it to be true</p> <p><math>s_4</math>: make an utterance only if he <i>wants to give the impression that he</i> wants to give the impression that he believes it to be true</p> <ul style="list-style-type: none"> <li>• (to continue, insert the phrase</li> <li>• italicized above each time)</li> <li>•</li> </ul> <p><math>s_n</math>: make an utterance only if he wants to give the impression that he . . . wants to give the impression that he believes it to be true</p>	<p><math>h_1</math> (reliance): respond to utterances by believing them to be true</p> <p><math>h_2</math> (trust): respond to utterances by coming to believe that the utterer believes them to be true</p> <p><math>h_3</math> (cageyness): respond to utterances by coming to believe that their utterer wants to give the impression that he believes them to be true</p> <p><math>h_4</math>: respond to utterances by coming to believe that their utterer <i>wants to give the impression that he</i> wants to give the impression that he believes them to be true</p> <ul style="list-style-type: none"> <li>• (to continue, insert the phrase</li> <li>• italicized above each time)</li> <li>•</li> </ul> <p><math>h_n</math>: respond to utterances by coming to believe that their utterer wants to give the impression that he . . . wants to give the impression that he believes it to be true</p>

Fig. 1

$sh_{i-1}$  could explain it *more completely*, as fewer additional assumptions needed to be made (see § II).

To illustrate this, consider again Case B, where Romeo utters a sentence he believes to be true, and Juliet responds by trusting that he believes it to be true. We can explain this piece of communication by saying that Romeo expected Juliet to trust him and that Juliet expected Romeo to be sincere, assuming a convention of sincerity and trust. We can also explain this by saying that Romeo expected Juliet to respond by recognizing his intention to give the impression that he believed the sentence to be true, and by saying that Juliet believed Romeo to have uttered the sentence because he wanted to give that impression. But this latter explanation is not yet complete. The coordination achieved by conformity to cageyness does not yet explain why Romeo was also sincere and why Juliet also trusted him. To explain this,

we need to say something in addition, for example that Juliet also believed that Romeo didn't want to give her a false impression. In this sense, the convention of cageyness provides a *less complete* explanation. In the same sense, any convention  $sh_i$  provides a less complete explanation than a convention  $sh_{i-1}$  for those cases of communication that can be explained by both. In other words, the larger the  $i$  of an  $sh_i$ , the more complicated the intentions and expectations we need to attribute to speakers and hearers; and the larger the  $i$  of an  $sh_i$ , the fewer violations of  $sh_i$ .

Therefore it might appear that the best candidate for linguistic conventions is to be found somewhere in the middle of the series  $\langle sh_1, sh_2, \dots, sh_n \rangle$ . But which one should we choose? The answer to this is that we should *not* choose one and ignore the rest. For those with higher indices than the one chosen will provide explanations in exceptional cases, and will not be violated. The ones with lower indexes than the one chosen will provide explanations without attributing excessively complicated intentions and expectations (in those cases where they apply). Rather, we should view linguistic conventions as complex disjunctive regularities. My suggestion is this:

A population  $P$  uses a language  $L$  iff there prevails in  $P$  the convention of conforming to  $s_1$  in  $L$  or  $s_2$  in  $L$  or . . . or  $s_n$  in  $L$  in uttering sentences of  $L$  and of conforming to  $h_1$  in  $L$  or  $h_2$  in  $L$  or . . . or  $h_n$  in  $L$  in responding to utterances of sentences of  $L$ .

Let us call such a convention a 'disjunctive convention in  $L$ '. Cases like B, C and more complex cases of deception can be viewed as instances of this regularity.

It might be objected that the complexity of disjunctive conventions prevents them from having psychological reality. But the complexity of the knowledge we need to attribute to speakers according to my suggestion does not constitute a disadvantage compared with Lewis's original account, as the common knowledge required for a convention is merely potential knowledge. Clause (6) of Lewis's definition of convention (see I) requires merely that the regularity, and the preference structure, be known *potentially*, i.e. that they would be known 'if one bothered to think hard enough'. Thus, the prevalence of a convention explains the conformity of participants by revealing the reasons they have for conformity without thereby assuming that they explicitly think of these reasons.

However, the objection from psychological reality might be taken further: speakers do not even have potential knowledge of a disjunctive regularity of the sort I have proposed, the opponent continues. For even if one 'bothered to think hard enough', one couldn't know that a regularity of, say, 50th order of complexity is being conformed to. We cannot even imagine a communicative situation in which a convention  $sh_{50}$  is being conformed to, but the corresponding convention  $sh_{49}$  is not.

Reply: perhaps it is in some sense correct to say that no such situation can be imagined, and that conformity to  $sh_{50}$  cannot even potentially be known.<sup>12</sup> But this is no objection to my proposal: in order to know that the regularity which constitutes a disjunctive convention is being conformed to one does not need to know that  $sh_{50}$  is being conformed to. All one needs to know (potentially) is that  $s_1$  or  $s_2$  or . . . or  $s_n$  is conformed to, and that  $h_1$  or  $h_2$  or . . . or  $h_n$  is conformed to. Knowledge that one of the disjuncts is true suffices for knowledge of the disjunction.

But, the opponent might continue, if knowledge of the disjunction that  $s_1$  is conformed to or  $s_2$  is conformed to or . . . or  $s_n$  is conformed to is never achieved through knowledge of the 50th disjunct, why include that disjunct at all? If no-one ever knows nor believes that a 50th disjunct is being conformed to, then no such knowledge or belief can be invoked to explain any utterance.

The reply highlights the *point* of making the disjunctions indefinitely long: it is plausible to assume that *if* we can comprehend conformity to a regularity  $s_n$ , then we can also comprehend conformity to  $s_{n+1}$  – there is no sharp cut-off point, no order of complexity  $n$  such that  $n$  is graspable and  $n+1$  is not. Given this assumption, we *cannot* let the series of disjuncts terminate. For suppose it were to terminate at some  $s_n$  (say at  $s_{50}$ ). Then we might as well have terminated at  $s_{n-1}$ , because if  $s_{n+1}$  is too complex, then so is  $s_n$ , according to our assumption. By sorites-type reasoning, we would soon end up saying that even  $s_1$  is too complex. Thus, the two disjunctions in a disjunctive convention need to be indefinitely long. Even though we know that many disjuncts are too complex to play a role in our reasoning, we do not know where exactly the cut-off point lies.

#### IV. Concluding Remarks

Why is my modified account to be preferred over Lewis's? – Let me stress that I do not deny that there are convention-like regularities of 'truthfulness and trust', at least if they are interpreted as regularities of sincerity and trust in my sense of § II. However, these regularities are frequently violated and therefore do not qualify as conventions in Lewis's sense. In order fully to account for the stability of such regularities, we must either assume the existence of additional regularities in the series, or claim that there prevails a disjunctive convention of the sort described above. For there are too many exceptions to regularities of this sort, and exceptions weaken the reasons agents have for conformity.

The advantages of my proposal can also be brought to bear on some well-known problems with Grice's account of speaker-meaning.<sup>13</sup> Lewis's account of linguistic convention can be seen as incorporating Grice's account of speaker-meaning (cf. his 1969, p. 155). Someone  $U$  who conforms to a

Lewisian convention of truthfulness and trust in a language  $L$  in uttering a sentence  $s$  of  $L$ , addressing an audience  $A$ , *also* speaker-means that  $s$  is true in  $L$ . He does so in the sense that he (1) intends  $A$  to come to believe that  $s$  is true; (2) intends  $A$  to recognize that  $U$  does (1); and (3) intends  $A$  to fulfil (1) on the basis of his fulfilling (2).

However, Grice came to change his original account, because of certain counterexamples: if a pupil, in answering an exam question, utters 'Waterloo was fought 1815', then he *says* that Waterloo was fought 1815, but does not speaker-mean it, since he obviously does not intend his teacher to come to believe that Waterloo was fought 1815, since he assumes that his teacher already knows that. Grice's response was to change clause (1), by requiring  $U$  to intend  $A$  to think that  $U$  believes that  $s$  is true, rather than intending  $A$  to believe that  $s$  is true (Grice 1968, p. 59).

Grice's revision corresponds to our convention of cageyness: an utterer who conforms to such a convention will also speaker-mean something in Grice's later sense. However, even the revised account is open to further counterexamples. There are cases of people saying that  $p$ , without thereby intending their audience to come to believe that the utterer himself believes that  $p$  (an example that comes to mind is an utterance of 'That's a rip-off' when haggling). Here too, the impression is that further Gricean revisions will lead to further counterexamples – in effect there will be an indefinite series of Gricean revisions that correspond to my series  $\langle sh_1, sh_2, \dots, sh_n \rangle$  of linguistic conventions.

This suggests a revision of Grice's notion of speaker-meaning, that is immune to this type of counterexample, which makes use of the disjunctive conventions I introduced above. I shall call it a notion of 'saying':

$S$  says (or speaker-means) that  $p$ , iff there is a language  $L$  and a sentence  $s$  of  $L$ , such that

- (a)  $s$  is true in  $L$  iff  $p$  (and this constitutes  $s$ 's truth condition),
- (b)  $S$  is member of a population among which a convention prevails of conforming to  $s_1$  in  $L$  or  $s_2$  in  $L$  or . . . or  $s_n$  in  $L$  and of conforming to  $h_1$  in  $L$  or  $h_2$  in  $L$  or . . . or  $h_n$  in  $L$ , and
- (c) in uttering  $s$ ,  $S$  (intentionally) conforms to the convention mentioned in (b).

The fact that my supplementation of Lewis's account yields this revision of Grice's account lends, I think, additional support to it, in so far as Lewis's account originally incorporated Grice's account of speaker-meaning.

Let me end on a note of caution. I showed that Lewis's explication of the ALR suffers from counterexamples. For example, Italians do not use Italian in Lewis's sense of 'using a language', because, given the frequency of lying,

answers to exam questions, haggling, etc., they violate the regularity of truthfulness and trust too often. My proposal for an emended account of the ALR avoids *this* problem and allows us to view lying, haggling, etc., as in conformity with linguistic convention. Nevertheless, there may be different kinds of communicative situation that still present a problem for the emended account. Fictional, ironical, and metaphorical utterances come to mind: it is not clear that a fictional utterance of ‘Romeo is a Montague’ is in conformity with a disjunctive convention concerning English; nor is this clear for ironical or metaphorical utterances. There may or may not be ways of dealing with these further difficulties within a Lewisian conventional framework. But that is a topic for another paper.<sup>14</sup>

## NOTES

1 Cf. Gilbert 1983, p. 388.

2 See Lewis 1975, p. 172.

3 Given that interest in restoring the connection is predominant.

4 This is a slightly abridged version of Lewis’s definition in his (1975), p. 165.

5 If we want to accommodate indexical languages, such as Lovetalk, we must speak instead of, for example, a function from sentences to Kaplanian ‘characters’ (which are in turn functions from contexts of utterance to truth conditions), or of a function from pairs of sentences and contexts of utterance to truth conditions.

6 For example, he can explain why Romeo emitted the sound ‘ti amo’: because (1) he wanted to let Juliet know that he loved her and (2) believed that he could achieve this by producing this sound. His belief was justified, because he knew that he and his addressee were speakers of Italian, and to be a speaker of Italian is to be a member of a population among which a convention of truthfulness and trust in Italian prevails. This, in turn, gave him reason to expect that he could achieve his aim by emitting the sound.

7 Someone might object as follows: In so far as anything is communicated in Case B, it is not a paradigm case of communication, for what is *communicated* is not what is *said*. We only need to account for the fact that people can say things with certain words, and we will thereby account for paradigm cases of communication (where what is said is what is communicated). Any non-paradigm case of communication ought then to be explained by recourse to what is said by the expression used in that case.

My reply is twofold: first, whether or not Case B is paradigmatic, I will show shortly that the fact alone that situations like Case B are frequent brings about counterexamples to Lewis’s explication of the ALR. Secondly, even those who pursue the strategy of accounting for non-paradigmatic cases of communication in terms of their notion of paradigmatic communication will admit that there is a sense in which linguistic communication is achieved in situations like Case B (and Case C below). McDowell (1980, pp. 128–31) does indeed admit this.

8 Probably there are convention-like regularities of reliability and reliance in actual languages that are frequently violated (and therefore do not count as conventions) but that are nevertheless stable. But this stability will be a side-effect of the prevalence of linguistic conventions of the sort I will introduce in § 3.

9 Someone might object that we must take Lewis’s words literally: conformity to truthfulness and trust consists just in *trying* to make only true utterances and *tending* to believe in the truth of utterances heard. But this suggestion is just as problematic as the reliability and reliance suggestion, for speakers frequently do not try to make only true utterances, and hearers frequently do not tend to believe what they are told. Therefore it is not the case that users of a language *L* generally have a convention of truthfulness and trust in *L* (literally understood), and the explication of the ALR fails.

- 10 See note 7 above again.
- 11 I am, for simplicity, only considering conventions that are combinations of speaker and hearer regularities with the same index number (see Figure 1). Of course, different combinations are also possible. This possibility does not affect the conclusion I am going to draw.
- 12 Even though the fact that my opponent and the reader understand me when I talk about  $sh_{50}$  shows that such knowledge would be possible in *some other* sense.
- 13 Cf. Grice (1968). McDowell, in his (1980), rejects Grice's notion because of these problems and offers as an alternative a notion of *saying* that is based on the idea that the primary function of assertion is the transmission of knowledge. See also Rumfitt (1995) for a good discussion of counterexamples to Grice's account.
- 14 I would like to thank the following for comments and suggestions: Brom Anderson, Dirk Boeckx, Mario Gomez-Torrente, Keith Hossack, Martin Kretschmer, Fraser MacBride, Karla Otero, David Papineau, Mark Sainsbury, Ruth Weintraub, and the audiences of presentations of earlier versions at Cumberland Lodge and the University of London.

## REFERENCES

- Gilbert, Margaret 1981. 'Game Theory and Convention.' *Synthese* 46, 41–93.
- Gilbert, Margaret 1983. 'Agreements, Conventions, and Language.' *Synthese* 54, 375–407.
- Grice, Paul 1968. 'Utterer's Meaning, Sentence Meaning and Word Meaning.' *Foundations of Language* 4, 1–18. Reprinted in John Searle (ed.), *The Philosophy of Language*. Oxford: Oxford University Press 1971, 54–70. Page references are to the reprinted version.
- Lewis, David 1969. *Convention*. Oxford: Basil Blackwell.
- Lewis, David 1975. 'Languages and Language', in Keith Gunderson (ed.), *Minnesota Studies in the Philosophy of Science*, vol. VII. Minneapolis: University of Minnesota Press, 3–35. Reprinted in David Lewis, *Philosophical Papers*, vol. I. Oxford: Oxford University Press, 163–88. Page references to the reprinted version.
- McDowell, John 1980. 'Meaning, Communication and Knowledge', in Zak Van Straaten (ed.), *Philosophical Subjects: Essays Presented to P. F. Strawson*. Oxford: Clarendon Press.
- Rumfitt, Ian 1995. 'Truth Conditions and Communication.' *Mind* 104, 827–62.
- Stenius, Erik 1967. 'Mood and Language Game.' *Synthese* 17, 254–74.

Received 3 February 1998

Max Kölbel, Instituto de Investigaciones Filosóficas, Universidad Nacional Autónoma de México, Ciudad Universitaria, 04510 México D.F.