

Pragmatics and epistemic vigilance: The deployment of sophisticated interpretative strategies

Diana Mazarella

This is the peer-reviewed version of the following article: Mazarella, D. (2015) Pragmatics and epistemic vigilance: The deployment of sophisticated interpretative strategies, *Croatian Journal of Philosophy*, XV(44), 183-199.

Abstract Sperber (1994) suggests that competent hearers can deploy sophisticated interpretative strategies in order to cope with deliberate deception or to avoid misunderstandings due to speaker's incompetence. This paper investigates the cognitive underpinnings of sophisticated interpretative strategies and suggests that they emerge from the interaction between a relevance-guided comprehension procedure and epistemic vigilance mechanisms. My proposal sheds a new light on the relationship between comprehension and epistemic assessment. While epistemic vigilance mechanisms are typically assumed to assess the believability of the output of the comprehension system (Sperber et al, 2010), I argue that epistemic assessment plays an additional role in determining this very output.

1. Communication and speaker's intentions

Current research on linguistic communication is grounded on the well-established assumption that “[h]uman communication is characterised, among other things, by the fact that communicators have two distinct goals: to be understood, and to make their audience think or act according to what is to be understood” (Sperber et al., 2010, p. 364). This is captured by Sperber and Wilson's (1986/1995) definition of ostensive-inferential communication. According to them, ostensive-inferential communication takes place when communicators produce an utterance (or any other ostensive stimulus) with the following two intentions: (i) an *informative intention* to inform the audience of something, and (ii) a *communicative intention* to inform the audience of one's informative intention (Sperber & Wilson, 1986/1995, p. 29).¹ More technically, the *informative intention* is defined as the intention to make manifest or more manifest to the audience an array of propositions I, whereas the *communicative intention* is defined as the intention to make it mutually manifest to audience and communicator that the communicator has this informative intention.²

Understanding requires the fulfilment of the communicative intention, that is, the recognition of the speaker's basic level/informative intention, but it might not require the fulfilment of the latter. This is the case with assertions. Understanding is achieved when it becomes mutually manifest to the communicator and the audience that the speaker has asserted that P. However, this does not require the assumption P to increase its manifestness, that is, to become more likely to be entertained and *accepted as true*. If we limit ourselves to assertions, we could say that the informative intention is about getting the audience to believe, and the communicative intention is about getting them to understand. Understanding is a matter of recognising what the speaker intends you to believe.

Crucially, an audience can understand an utterance without believing what they have understood. In this case, the communicative intention is fulfilled without the corresponding informative intention being fulfilled. As Wilson and Sperber (2004, p. 611) suggest, typically “[w]hether the informative intention itself is fulfilled depends on how much the audience trusts the communicator”. While understanding is underpinned by the pragmatic ability to recover the speaker's communicated meaning on the basis of linguistic and contextual cues, the epistemic assessment which leads to its acceptance (or rejection) is carried out by what Sperber et al. (2010) call ‘epistemic vigilance’.

¹ In contrast with Grice (1957), Sperber and Wilson reject the idea that the communicator must have a third-level intention that the addressee's recognition of her informative intention should be at least part of his reason for fulfilling this. This allows Sperber and Wilson's account of ostensive-inferential communication to cover the whole continuum from pure cases of ‘showing’ to pure cases of ‘meaning’. See also Sperber and Wilson (2015).

² See Sperber and Wilson (2015) for a detailed discussion on the notion of ‘manifestness’.

Before exploring the nature of epistemic vigilance mechanisms, I intend to show that the possibility of comprehension without acceptance/belief is warranted by the definition of communicative and informative intentions provided by Sperber and Wilson. To begin with, it is worth noting that their definition of manifestness is epistemic in nature: a proposition is manifest to an individual at a given time to the extent that he is likely to some positive degree to entertain it and *accept it as true*. Given that communicative and informative intentions are conceived of as intentions to change the degree of manifestness of certain assumptions, the following question seems to arise. Is the rejection of the communicated content compatible with the idea that manifestness involves acceptance on the interpreter's part?³ In order to answer this question, it is worth stressing that in those circumstances in which the addressee understands a piece of communicated information without ending up believing it, the communicative intention is fulfilled without the corresponding informative intention being fulfilled. The informative intention is *recognised* (rather than fulfilled). The fulfilment of the communicative intention entails that the fact that the communicator has a certain informative intention, let's call it " i_i ", is made mutually manifest. That is, the fact that the communicator has the intention i_i is likely to be accepted as true by the audience (and the communicator) and this is itself manifest. This does not seem to be problematic: while the audience is likely to accept as true the fact that the speaker has the intention i_i (i.e. the intention to make manifest or more manifest to the audience an array of propositions I_i), the audience is not bound to accept as true any of the propositions which are included in I_i . The audience accepting as true the array I_i would correspond to the fulfilment of the informative intention (but it is not a condition for its recognition).

Importantly, the output of the pragmatic system is metarepresentational in nature. According to Relevance Theory, utterance interpretation is a process that starts with a metarepresentation of an attributed utterance ('The speaker uttered u ') and ends with a metarepresentation of an attributed thought or set of thoughts ('The speaker communicated I '). Understanding requires entertaining and accepting as true this metarepresentational output.

2. Epistemic vigilance

Sperber et al. (2010) suggest that humans have developed "a suite of cognitive mechanisms", which is targeted at the risk of misinformation in communication. Each of them is likely to be specialised in one of the many kinds of considerations relevant to warranting (or undermining) epistemic trust.

But what exactly is 'epistemic trust'? It can be defined as the willingness to believe the communicator and accept her claims as true. Communicators are not always competent or benevolent and communication is thus open to the risk of misinformation. A competent communicator possesses genuine information (rather than misinformation or no information), whereas a benevolent communicator is willing to share the information he has (as opposed to asserting false information because of indifference or malevolence). If communication has to remain advantageous on average (as its pervasiveness in our social interactions suggests it is), humans have to deploy an ability to calibrate their epistemic trust. This ability is 'epistemic vigilance'.

³ Thanks to Steve Oswald for raising this question at the *PragLab Research Colloquium in Linguistics* in Fribourg (May 2015).

Sperber et al. (2010) conceive of epistemic vigilance as a cognitive adaptation for social exchange. As Cosmides and Tooby (1992, p. 166) suggest, “each cognitive specialisation is expected to contain design features targeted to mesh with the recurrent structure of its characteristic problem type”. Thus, a closer investigation of its ‘problem type’ will shed some light on the nature and function of the cognitive mechanisms underpinning epistemic vigilance as a whole.⁴

The ‘problem type’ that represents the target of epistemic vigilance is the risk of misinformation in communication. Misinformation can be either accidental or intentional. The former is often the result of speaker’s incompetence, the latter of speaker’s malevolence. An incompetent speaker may communicate information that is false because she takes it to be true; a malevolent speaker may communicate false information with the intention of deceiving her interlocutor.

These alternative and recurrent features of misinformation suggest that some of the epistemic vigilance mechanisms should check for the reliability of the source of information, where reliability is a function of both speaker’s competence and speaker’s benevolence. In other terms, epistemic vigilance should help us with monitoring *who* to believe (i.e. individuals who are both competent and trustworthy).

A growing body of research on the development of the epistemic vigilance capacity towards the source indicates that this ability emerges very early in development (for a review, see e.g. Harris (2012), Robinson and Einav (2014)). Some form of epistemic vigilance may be present from the very age infants have actually been tested. For instance, as reported by Koenig and Harris (2007), when 16-month-olds saw pictures of familiar objects and heard accurate/inaccurate labels from (a) a human looking at the picture, (b) a human with her back to the picture, (c) an audio speaker, they tended to be surprised when label (a) was false, when label (b) was true, and not surprised either way by (c).

By 2 to 4 years of age, children employ a number of criteria for evaluating the reliability of the speaker. They show selective trust based on past accuracy (2-year-olds, see Koenig and Harris (2007)), speaker’s attitude (indications of certainty/hesitation) (3 year-olds, see Matsui, Rakoczy, Miura and Tomasello, (2009)), true knowledge vs. past accuracy (4-year-olds, see Einav and Robinson (2011)).

As far as epistemic vigilance towards deception is concerned, children become able to cope with intentional deception from 4 to 6 years of age (Mascaro & Sperber, 2009). This capacity requires sophisticated mindreading abilities, as the interpreter needs to combine a first-order attribution of belief (‘The speaker believes that not-P’) with a second-order attribution of intention (‘The speaker wants me to believe that P’).

The reliability of the source of information, however, is not the only factor affecting the believability of a piece of communicated information. The content of information may itself be more or less believable, independently of its source (with tautologies and logical contradictions lying at the two extremes of a continuum of believability). Thus, Sperber et al. (2010) argue for the existence of a second cluster of epistemic vigilance mechanisms, that is, mechanisms which assess the quality of the incoming information (i.e., *what* to believe). They check its factual plausibility by assessing its consistency with existing knowledge and its degree of evidence. According to Sperber et al. (2010), the beliefs against which the communicated information is tested are those that are automatically activated by the comprehension process and used in the pursuit of relevance. These are a subset of the mental

⁴ Both Sperber and Cosmides and Tooby advocate the massive modularity view of the mind, that is, the view that the mind is a system of evolved cognitive mechanisms that are dedicated to a particular task (hence domain-specific) and interact with each other in constrained ways.

encyclopaedia of the addressee, and provide the ground for an “imperfect but cost-effective epistemic assessment (Sperber et al. 2010, p. 374). When the result of this assessment is a contradiction, there are three possible outcomes: (i) if the source is taken as trustworthy and the background beliefs of the interpreter that conflict with the incoming information are not held with much conviction, these beliefs are corrected; (ii) if the source is not regarded as trustworthy, the new information is rejected; (iii) if the source is regarded as authoritative and the conflicting background beliefs are held confidently, some process of (typically conscious) coherence checking is triggered. The choice among (i), (ii), and (iii) partly depends upon the output of epistemic vigilance mechanisms focused on the source (the speaker).

In conclusion, according to Sperber et al. (2010), the gap between comprehension and acceptance/belief is bridged by epistemic vigilance mechanisms, which play a significant role in filtering incoming information with the aim of minimising the risk of misinformation. In what follows, I will address the question of what role (if any) epistemic vigilance mechanisms may play in the comprehension process itself.

3. Comprehension and epistemic assessment

3.1 Sperber et al.’s (2010) ‘pragmatic’ model

Sperber et al. (2010) have recently proposed a model of the relation between comprehension and epistemic assessment that, for reasons which will soon become apparent, we shall call the ‘pragmatic’ model. According to Sperber and colleagues, comprehension and epistemic assessment are *parallel* processes which are triggered by the very same act of ostensive communication. While comprehension is underpinned by a relevance-guided comprehension procedure, epistemic assessment is carried out by dedicated mechanisms which contribute to the capacity for ‘epistemic vigilance’.

Relevance Theory claims that comprehension is driven by the expectations of relevance which are raised by every ostensive stimulus. Specifically, every ostensive stimulus conveys a presumption of its own optimal relevance, that is, the expectation that the stimulus will be relevant enough to the addressee (to be worth processing) and that it is the most relevant one compatible with the speaker’s abilities and preferences. This presumption justifies the adoption of the following comprehension heuristic:

(1) Relevance-guided comprehension procedure

- a. Follow a path of least effort in computing cognitive effects: Test interpretative hypotheses (disambiguations, reference resolutions, implicatures, etc.) in order of accessibility.
- b. Stop when your expectations of relevance are satisfied.

(Wilson & Sperber, 2004, p. 613)

Interestingly, Sperber et al. (2010) maintain that comprehension and epistemic assessment are interconnected aspects of a single process whose goal is to make the best of communicated information and they suggest that considerations of believability play a crucial role in the comprehension process itself:

“We claim that, whether he ends up accepting it or not, the hearer interprets the speakers as asserting a proposition that would be relevant enough to him provided that he accepted it” (Sperber et al., 2010, p. 386)

This is, however, a ‘hypothetical’ role: comprehension initially proceeds *as if* the interpretative hypotheses under construction were to be accepted as true. Given such an assumption, the first hypothesis that satisfies the addressee’s expectations of relevance is attributed to the speaker as her intended meaning. No *actual* assessment of the believability of that hypothesis needs to take place at this stage. Importantly, the presumption of optimal relevance which is communicated by every ostensive stimulus need not to be true or accepted as true: the speaker might fail to achieve relevance either because of incompetence or malevolence (Sperber & Wilson, 1986/1995, pp. 158-159). However, the very fact that the presumption is communicated is enough to guide the interpretative process. To illustrate this point, Sperber and colleagues discuss the following example. Andy and Barbara have decided to throw a party and Barbara has asked Joan to bring a bottle of champagne.

- (2) a. *Andy (to Barbara):* A bottle of champagne? But champagne is expensive!
- b. *Barbara:* Joan has money.

Imagine that Andy had previously assumed that Joan was a junior underpaid academic. In the context at issue, Barbara’s utterance would make a relevant contribution to the discussion if Andy interpreted it as communicating that *Joan has enough money to be easily able to afford champagne*. The interpretation that *Joan has some money* (which is not only true but also compatible with Andy’s background belief) would make little sense as a conversational move at this point of the conversation. Considerations of relevance lead Andy to interpret Barbara’s utterance in the expected way: the interpretation that *Joan has enough money to be easily able to afford champagne* is relevant to Andy provided that he accepts it as true. Andy may decide to reject the communicated information (for instance, because he thinks that Barbara does not know that Joan is only a teaching assistant who is paid on an hourly basis) but, whether or not he ends up believing it, he will interpret it in order to optimise its (intended) relevance.

In line with this, Origgi (2008) suggests that interpretation involves a ‘stance of trust’ that our interlocutors will provide relevant information for us. Any departure from the satisfaction of our expectations of relevance may result in a revision or a withdrawal of the initial trust with which we approach the interpretative process: this is why the stance of trust is ‘dynamic’ – it is only tentative and labile, but it plays a crucial role in determining the output of the comprehension system.

It is crucial to notice, though, that while Sperber et al. (2010) see comprehension and epistemic assessment as parallel processes triggered by the same piece of communicative behaviour (‘The speaker has uttered *u*’), they think of their interaction as limited in scope. The only role of the epistemic vigilance system is to assess the believability of the interpretation *resulting from* the comprehension process (in light of considerations about both communicator’s reliability and content’s believability). In what follows, I suggest that the interaction between comprehension and epistemic assessment has a wider scope than previously assumed and, as a result, it may be more finely articulated.

3.2 Competence, benevolence and interpretative strategies

It is worth beginning our investigation of the relationship between comprehension and epistemic assessment by considering examples that give rise to clear off-line intuitions about how our interpretative practice might be affected by considerations about the moral and epistemic trustworthiness of our interlocutors.

Sperber (1994) invites us to consider the following scenario. Imagine that Carol and John are going to a party and they have planned to leave their child at home with the baby sitter. The baby sitter usually leaves at midnight. That day, however, thinking that the party would be great fun, Carol has made a special arrangement with the babysitter and she will stay until one. Crucially, Carol does not know that John is aware of this. Later that night, Carol is not enjoying the party and, at around 11.30pm, she says to John “It’s late” expecting him to think that it is time to go home because of the baby sitter. The interpretation of Carol’s utterance “It’s late” depends on whether the addressee (John) *trusts* the speaker (specifically, on whether he assumes that Carol is behaving benevolently). If John assumes Carol’s benevolence, he will be bound to misunderstand her. She could not intend to communicate something that she knows to be false, so he would take it that she intended to communicate that it is late with respect to some other schedule or expectation (for instance, that it is late if they want to catch the last train to get home). On the other hand, if he recognises that Carol is trying to deceive him, he will correctly attribute to her the intention to communicate that it is time to go home because of the baby sitter.

Let us consider a modified version of this scenario. Imagine that Carol and John are at the party and no special arrangement has been made with the baby sitter, who will leave, as usual, at midnight. Unbeknownst to Carol, John is very worried about a delivery that should have been made that very day. At 11.30pm, Carol says to John “It’s late” expecting him to think that it is late because of the baby sitter. Because he is caught up in his thoughts, the first interpretation to come to John’s mind is that the delivery is late. Once again, the interpretation of the utterance “It’s late”, which is eventually attributed to Carol, depends on whether the addressee trusts the speaker (this time, on whether John assumes that Carol is competent, that is, she possesses genuine information as opposed to misinformation or no information). If he does, he will be bound to misunderstand her by mistakenly attributing to her the intention to communicate that the delivery is late. If he realises that she could not intend to communicate something that she does not know, he would take it that she intended to communicate that it is time to go home because of the baby sitter.

These examples, taken together, suggest that consideration of the speaker’s benevolence and competence may affect the way in which we interpret what she says. An adequate account of pragmatic processing should shed some light on how this happens to be the case. With regard to this, Sperber (1994) suggested that competent interpreters have sophisticated interpretative strategies at their disposal, which allow them to cope with deliberate deception or to avoid misunderstandings due to speaker’s incompetence. This proposal has been given only a relatively marginal role within the development of Relevance Theory in the following years. In what follows, I present the interpretative strategies proposed by Sperber (1994) and suggest the existence of an interesting link between these strategies and the operations of epistemic vigilance mechanisms.

Sperber (1994) suggests the existence of three interpretative strategies, which he labels ‘naïve optimism’, ‘cautious optimism’ and ‘sophisticated understanding’, which can be seen as different versions of the relevance-guided comprehension procedure:

(1) **Relevance-guided comprehension procedure**

- a. Follow a path of least effort in computing cognitive effects: Test interpretative hypotheses (disambiguations, reference resolutions, implicatures, etc.) in order of accessibility.

- b. Stop when your expectations of relevance are satisfied.

As emphasised by Wilson and Sperber, clause (b) of the relevance-guided comprehension procedure “[..] allows for varying degrees of sophistication in the expectations of relevance with which an utterance is approached.” (Wilson & Sperber, 2004, p. 625). Importantly, the difference between the three strategies relies on different assumptions about the communicator’s competence and benevolence, which in turn raise different expectations of relevance (hence determine different stopping points in interpretation). A naïvely optimistic hearer takes for granted that the communicator is behaving both benevolently and competently: he takes the communicator to be competent enough to avoid misunderstanding, and benevolent enough not to lead him astray. Thus he expects ‘actual optimal relevance’. A naïvely optimistic hearer looks for an interpretation that is relevant enough to him and he assumes that it is the intended one. In contrast, a cautiously optimistic interpreter assumes the communicator to be benevolent, but not necessarily competent: the communicator, in fact, may not know what is in the addressee's mind and thus fail to produce the most relevant stimulus for him. As a consequence, he looks for ‘attempted optimal relevance’. Finally, a sophisticated interpreter drops not only the assumption that the communicator is behaving competently, but also that she is behaving benevolently. Then the expectations of relevance that guide the comprehension procedure and determine its stopping point are expectations of ‘purported optimal relevance’. The following table illustrates the three different versions of the relevance-guided comprehension procedure (which differ with regard to clause (b)):

Three versions of the relevance-guided comprehension procedure:

- (a) Follow a path of least effort in computing cognitive effects:
Test interpretative hypotheses in order of accessibility.

<i>Naïve optimism</i>	(b ₁)	Stop when your expectations of <i>actual optimal relevance</i> are satisfied (i.e. stop at the first relevant enough interpretation)
<i>Cautious optimism</i>	(b ₂)	Stop when your expectations of <i>attempted optimal relevance</i> are satisfied (i.e. stop at the first interpretation that the communicator <i>might have thought</i> would be relevant enough to you)
<i>Sophisticated understanding</i>	(b ₃)	Stop when your expectations of <i>purported optimal relevance</i> are satisfied (i.e. stop at the first interpretation that the communicator <i>might have thought would seem</i> relevant enough to you)

Successful interpretative paths may require the adoption of sophisticated interpretative strategies. Let us consider, for instance, the examples discussed above. With regard to the delivery-example, if John were a naïvely optimistic interpreter, he would attribute to Carol the first interpretative hypothesis that is relevant enough to him. The first interpretation that comes to John’s mind is that the delivery is late. Given its relevance to John, a naïve interpreter would retain it and mistakenly attribute it to the speaker. But what if John adopted the cautiously optimistic version of the relevance-guided comprehension procedure? John would not take for granted Carol’s competence and he would be vigilant to the possibility that Carol may not know what he knows (and may consequently fail in her attempt to make the relevant information that she intends to convey more accessible than any other possible interpretation). John would realise that

Carol could not have intended the interpretative hypothesis *The delivery is late* to occur to him (precisely because she does not know that he is waiting for a delivery). Carol could not have thought that this interpretation would be relevant enough to him as she, Carol, has no thoughts of any sort involving this delivery. Thus, the comprehension procedure would go further and test the next most accessible interpretative hypothesis. For instance, it would access and assess the interpretation that it is time to go home because of the baby sitter. Since John takes it that Carol might have thought this interpretation to be relevant enough to him (as in fact it is), the interpretation is selected and attributed to Carol. With regard to the deceptive version of this example, it is possible to show how both a naïve and a cautiously optimistic interpreter would fail in attributing to the speaker the intended interpretation. Let us assume that the first interpretation to come to John's mind is the (intended) interpretation that it is time to go home because of the baby sitter. However, John immediately realises that this is not the case, as he knows that Carol and the baby sitter have made a special arrangement for that night and that the baby sitter will leave later than usual. If John were a naively optimistic interpreter, he would discard that interpretation, as it is not relevant to him (he knows it to be false). If he were a cautiously optimistic interpreter, he would also discard it given that it is not an interpretation that Carol might have thought would be relevant to him (having made the arrangement herself, Carol knows that it is not the case that it is time to go home because of the baby sitter). Only the adoption of a more sophisticated interpretative strategy would allow John to correctly attribute this interpretation to Carol. John would realise that Carol might have thought that this interpretation would *seem* relevant to him (as she does not know that John is aware of this special agreement) and, if he had reasons to think that she might want to deceive him, he would end up attributing that interpretation to her.

3.3 The role of epistemic vigilance in comprehension

Utterance interpretation may depend on considerations about the speaker's competence and/or benevolence. The issue of what brings such considerations to bear on the interpretative process, however, has not been addressed within the literature. My proposal is that the expectations of relevance which guide the comprehension procedure and determine its stopping point are directly modulated by the operations of epistemic vigilance mechanisms. That is, epistemic vigilance mechanisms can modulate the hearer's expectations of relevance (i.e. from 'actual' to 'attempted' or 'purported' optimal relevance) and assess whether the interpretative hypothesis under construction satisfies these expectations.

If the interpreter is vigilant towards the speaker's competence and finds reasons to doubt it, he will expect 'attempted' optimal relevance. As a consequence, he will stop at the first relevant interpretation that the speaker might have thought would be relevant to him (as described in the cautiously optimistic version of the relevance-guided comprehension procedure). If the interpreter is vigilant towards the speaker's competence as well as her benevolence and he realises that the speaker may want to deceive him, he will expect 'purported' optimal relevance. In this case he will stop at the first interpretation that the speaker might have thought would seem relevant enough to him.

In the same vein as my proposal, Padilla Cruz (2012) has suggested that epistemic vigilance should be considered as the trigger for a shift in interpretative strategies. For instance, if epistemic vigilance detects that the interlocutor is not a very competent communicator, it may trigger a shift from naïve optimism to cautious

optimism. I believe, though, that the recent work on epistemic vigilance should be seen as encompassing Sperber's (1994) original proposal. Once epistemic vigilance is brought into the picture, the three interpretative strategies are found to be redundant. For instance, a cautiously optimistic interpreter may be seen not as an interpreter who is prompted to adopt a particular strategy by his epistemic vigilance mechanisms (as Padilla Cruz suggests), but rather as an interpreter who is actively monitoring the speaker's competence through his epistemic vigilance mechanisms. A very interesting and plausible picture emerges: the three interpretative strategies described above may simply be an *epiphenomenon* of the interaction between a single comprehension procedure and epistemic vigilance mechanisms.

Before developing this proposal with regard to the examples under discussion, it is worth noting that Sperber and colleagues suggest that epistemic vigilance mechanisms towards the source of information can deliver either general impressions of trustworthiness or more costly assessments that result from context-sensitive evaluations of the reliability of the speaker. With regard to the latter, they note the following:

Clearly, the same informant may be competent on one topic but not on others, and benevolent towards one audience in certain circumstances, but not to another audience or in other circumstances. This suggests that trust should be allocated to informants depending on the topic, the audience, and the circumstances. (Sperber et al., 2010, p. 369)

The definition of trustworthiness provided by Sperber et al. (2010) is thus intrinsically context-dependent; and it could not be otherwise. To elaborate on this point, let us focus on speaker's competence. For every speaker, there is always some information that she does not possess and some false assumptions that she takes to be true. However, this is not what 'competence' is about. If this were the case, every speaker would have to be classified as incompetent and would not be entitled to receive our epistemic trust. Competence has a narrower and context-sensitive scope: the same communicator may be competent on one topic but not on others. The investigation of epistemic vigilance mechanisms that can assess competence (as well as benevolence) in a context-sensitive way will prove to be crucial for a general understanding of epistemic vigilance, and my proposed interaction with the comprehension system.

To clarify the dynamics of the hypothesised interaction between the relevance-guided comprehension procedure and epistemic vigilance mechanisms, let us consider again the example "It's late", where the first interpretative hypothesis to come to the hearer's mind is that the delivery is late but this does not correspond to the intended interpretation. My suggestion is that the construction of an interpretative hypothesis provides a hypothesised topic of conversation. This, in turn, serves as input to epistemic vigilance mechanisms which assess the competence of the speaker on a particular topic. In this case, the interpretative hypothesis that the delivery is late provides a hypothesised topic of conversation (i.e. the delivery) with regard to which epistemic vigilance mechanisms assess Carol's competence. These mechanisms access the piece of information that Carol does not know that John is waiting for a delivery. As a consequence, an incompatibility between the speaker's system of beliefs and the interpretative hypothesis under construction is detected. This inhibits the comprehension procedure and prompts it to access (and assess) the next most accessible interpretative hypothesis.

Let us now consider the deceptive version of this example, where Carol's utterance "It's late" is intended to be interpreted as implicitly communicating that it's time to go home because of the baby sitter. Suppose that the first interpretation to come to John's mind in this context is the interpretation that it is time to go home because of the baby sitter (where this conclusion is warranted by the explicature of "It is late" and the implicated premise *The baby sitter typically leaves at midnight*). Once again, this provides a hypothesised topic of conversation, that is, the baby-sitter arrangement, which, as in the previous example, triggers epistemic vigilance mechanisms which assess the speaker's competence on that topic. These mechanisms access the piece of information that Carol believes that it is not the case that the baby sitter will leave at midnight. Given that the interpretative hypothesis is not compatible with the speaker's epistemic state, a cautiously optimistic hearer would discard this hypothesis and look for a different interpretation. However, if the hearer's epistemic vigilance mechanisms targeted at assessing the speaker's benevolence detect that the speaker has the deceptive intention of making the hearer (falsely) believe that it is time to go home because of the baby sitter (and if this intention is also compatible with what the hearer believes the speaker takes to be the hearer's epistemic state, i.e. Carol does not know that John knows that the baby sitter will stay until one), the interpretation would be retained and attributed to the speaker. As Sperber and colleagues suggest, "[w]hen epistemic vigilance is targeted at the risk of deception, it requires an understanding not only of the communicator's epistemic states but also of her intentions, including intentions to induce false beliefs in her audience" (Sperber et al, 2010, p. 372). In our example, the hearer understands that Carol believes that not-P but she wants him to believe that P (P = it is time to go home because of the baby sitter).

As the discussion of these examples illustrates, not only do epistemic vigilance mechanisms affect the *believability* of a piece of communicated information (as proposed by Sperber and colleagues), but they also contribute to the assessment of the interpretative hypotheses under construction. That is, not only do they establish whether an interpretation attributed to the speaker (i.e. the output of the comprehension procedure) is allowed to enter the 'belief box' of the interpreter, but they also assess whether an interpretative hypothesis under construction is to be retained and attributed to the speaker as the intended interpretation (i.e. whether it ends up being the output of the comprehension system or not). As far as this latter role is concerned, epistemic vigilance mechanisms may filter out interpretative hypotheses that, although relevant, are incompatible with the speaker's mental states (i.e. her beliefs and desires). In this case, they may prompt the comprehension process to continue and assess further interpretative hypotheses. In other circumstances, they may prevent the comprehension procedure from abandoning an interpretative hypothesis that is irrelevant (to the hearer, e.g. he knows that it is false) but compatible with the speaker's mental states (e.g. her intention to induce false belief in the hearer).

3.4 Implications

It is worth investigating the implications of this proposal for the relationship between comprehension and epistemic assessment. Specifically, the question of whether the epistemic assessment that is involved in comprehension exhausts the validation-process for communicated information shall be addressed. According to my proposal, the effects of epistemic vigilance on comprehension and acceptance/belief can be either *simultaneous* (when no further epistemic assessment beyond that involved in comprehension is needed) or *serial*. When the addressee arrives at the intended

interpretation via the recognition that the communicator is trying to deceive him, it is plausible to assume that no further epistemic assessment will be undergone and the output of the comprehension process will be automatically prevented from entering the addressee's belief box. This seems to be the case in the deceptive version of the example discussed above. If John correctly reaches the intended interpretation that it is time to go home because of the baby sitter via the recognition of Carol's deceptive intention, he won't further assess the believability of this interpretation after attributing it to Carol. However, when the role played by epistemic vigilance is that of warranting an interpretation that is compatible with the speaker's system of beliefs, it is still an open issue whether or not the interpretation selected is worth being accepted as true (i.e. believed). For instance, in the delivery-example discussed above, John recognises that Carol could not intend to communicate that the delivery is late, as she is not aware that he is waiting for a delivery. As a consequence, he looks for an interpretation that is compatible with Carol's system of beliefs, e.g. that it is late if they want to catch the last train. Once this interpretation has been attributed to Carol as the intended interpretation, it is still an open question whether John should accept it as true or not. For instance, it might contradict some of his strongly held beliefs concerning the recent introduction of a 24-hour train service and could thus be rejected as false. In this case, rejection (or acceptance/belief) will be the result of a process of epistemic assessment that takes place on the *output* of the comprehension system.

Finally, let us explore the implications of this hypothesised interaction between the comprehension system and epistemic vigilance mechanisms with regard to the relationships among Sperber's (1994) three interpretative strategies:

Much of everyday communication takes place between people who are benevolent to one another and who know one another well enough. In such circumstances, cautious, and even naïve optimism can serve as default interpretation strategies [...] Still, *when the optimistic strategies fail*, a competent hearer resorts to the sophisticated strategy. (Sperber, 1994, p. 198, *my emphasis* (DM))

In line with this, if we assume that more sophisticated interpretative strategies are implemented by the operations of epistemic vigilance mechanisms, we can endorse the idea that optimistic interpretative strategies might represent a 'preferred' option, but reject the claim that optimistic strategies must fail in order for the interpreter to resort to 'sophisticated understanding'. Rather, addressees may start by being vigilant and only grant trust if they have no reasons to doubt the communicator's competence and benevolence. If they do have reasons to doubt her trustworthiness, they will downgrade their initial expectations of 'actual' optimal relevance to expectations of 'attempted' or 'purported' optimal relevance. As emphasised by Sperber et al. (2010), epistemic vigilance is not the opposite of trust, but it is the opposite of *blind* trust. To clarify this point, Sperber et al. (2010, p. 346) develop the following enlightening analogy: when we walk down the street through a crowd of people, we typically do not have any hesitation about walking among them, despite the risk of being accidentally or intentionally hit by them. However, we do monitor the trajectory of others and if we detect the presence of a careless or aggressive individual, we raise our level of vigilance. This low-level, unconscious, vigilance allows us to enjoy our stroll while preventing us from any risk of collusion. In other words, our mutual trust is buttressed by our mutual vigilance.

4. Conclusions

While the interaction between the comprehension system and epistemic vigilance mechanisms has not been much explored, its centrality has already been recognised:

[...] the abilities for overt intentional communication and epistemic vigilance must have evolved together, and must also develop together and *be put to use together*. (Sperber et al., 2010, p. 360, *my emphasis* (DM))

This passage suggests three different perspectives that are relevant to the investigation of epistemic vigilance in communication: an evolutionary perspective, a developmental perspective, and a 'pragmatic' perspective. The main focus of this paper was to explore the 'pragmatic' dimension of this interaction.

According to Sperber et al. (2010), comprehension and epistemic assessment are triggered by any act of ostensive communication: while comprehension follows a relevance guided procedure which selects the interpretative hypothesis that would be relevant to the hearer if he accepted it, epistemic vigilance mechanisms assess the speaker's reliability and the factual plausibility of the communicated content and establish whether the selected interpretation is indeed worth being accepted as true. Going beyond this proposal, I argued that, not only do epistemic vigilance mechanisms assess the believability of the output of the comprehension system, but they also play a crucial role in determining that output. As discussed above, epistemic vigilance mechanisms towards the source modulate the expectations of relevance that drive the comprehension procedure. If they detect that the speaker is not competent, they downgrade the initial expectation of actual optimal relevance to *attempted* optimal relevance. Similarly, if epistemic vigilance mechanisms towards deception detect that the speaker is not benevolent, they set the expectations of relevance at an even lower grade, that is, *purported* optimal relevance. As a consequence, epistemic vigilance mechanisms may filter out interpretative hypotheses that are incompatible with assumptions about the speaker's epistemic state or retain interpretative hypotheses that are irrelevant to the addressee (e.g. because he knows them to be false) but compatible with the speaker's deceptive intention. This has important consequences for the relationship between comprehension and epistemic assessment. Crucially, epistemic assessment taking place *after* comprehension can be made redundant whenever the selection of the intended interpretative hypothesis is grounded on the recognition that the speaker has a deceptive intention. If it is by realising that the speaker has the intention to mislead him that the addressee recognises what the speaker intended to communicate, there is no need to further assess the believability of the misleading pieces of communicated information. However, when the role of epistemic assessment in comprehension is that of warranting an interpretation that is compatible with the speaker's epistemic state, it remains an open question whether the communicated content is worth being accepted as true. In this latter circumstance, further epistemic assessment would typically follow comprehension.

Interestingly, this proposal has potential implications for the development of pragmatic abilities. Sperber's (1994) hypothesised that the three interpretative strategies corresponding to naïve optimism, cautious optimism and sophisticated understanding might represent different developmental stages. That is, children would start as naïve optimistic interpreters and subsequently deploy increasingly sophisticated

interpretative strategies, which are based on the recognition that the speaker may not be competent or benevolent. In light of my proposal on the interaction between the comprehension system and epistemic vigilance mechanisms, such strategies may emerge thanks to the unfolding of epistemic vigilance capacities in the child's cognitive development. This sheds a new light on the relation between communicative abilities and other types of metarepresentational capacity (in particular, epistemic vigilance), which is open to further empirical investigation.

Acknowledgements

Many thanks to Robyn Carston, Deirdre Wilson and Dan Sperber for invaluable discussions on the content of this paper. This work has also benefited from the insightful comments of the participants at the *Philosophy of Language and Linguistics Workshop* (Dubrovnik, September 2014).

References

- Cosmides, L., & Tooby, J. (1992). Cognitive adaptations for social exchange. In J. H. Barkow, L. Cosmides, & J. Tooby (Eds.), *The adapted mind: Evolutionary psychology and the generation of culture* (pp. 163–228). New York: Oxford University Press.
- Einav, S., & Robinson, E. J. (2011). When being right is not enough: Four-year-olds distinguish knowledgeable from merely accurate informants. *Psychological Science*, 22, 1250–1253.
- Grice, H. P. (1957). Meaning. *Philosophical Review*, 66, 377–388.
- Harris, P. L. (2012). *Trusting what you're told: How children learn from others*. Cambridge, MA: Harvard University Press.
- Koenig, M. a., & Harris, P. L. (2007). The Basis of Epistemic Trust: Reliable Testimony or Reliable Sources? *Episteme*, 4(03), 264–284.
- Mascaro, O., & Sperber, D. (2009). The moral, epistemic, and mindreading components of children's vigilance towards deception. *Cognition*, 112(3), 367–380.
- Matsui, T., Rakoczy, H., Miura, Y., & Tomasello, M. (2009). Understanding of speaker certainty and false-belief reasoning: A comparison of Japanese and German preschoolers. *Developmental Science*, 12(4), 602–613.
- Origi, G. (2008). Trust, authority and epistemic responsibility. *Theoria. Revista de Teoría, Historia Y Fundamentos de La Ciencia*, 23(61), 35–44.
- Padilla Cruz, M. (2012). Epistemic vigilance, cautious optimism and sophisticated understanding. *Research in Language*, 10(4), 365–386.
- Robinson, E. J., & Einav, S. (2014). *Trust and Skepticism: Children's Selective Learning from Testimony*. New York: Psychology Press.

- Sperber, D. (1994). Understanding verbal understanding. In J. Khalfa (Ed.), *What is intelligence?* (pp. 179–198). Cambridge: Cambridge University Press.
- Sperber, D., Clément, F., Heintz, C., Mascaro, O., Mercier, H., Origgi, G., & Wilson, D. (2010). Epistemic Vigilance. *Mind & Language*, 25(4), 359–393.
- Sperber, D., & Wilson, D. (1995). *Relevance: Communication and Cognition* (2nd ed.). Oxford: Blackwell Publishing.
- Sperber, D., & Wilson, D. (2015). Beyond speaker's meaning. *Croatian Journal of Philosophy*.
- Willis, J., & Todorov, A. (2006). First impressions: making up your mind after a 100-ms exposure to a face. *Psychological Science*, 17, 592–598.
- Wilson, D., & Sperber, D. (2004). Relevance Theory. In L. R. Horn & G. L. Ward (Eds.), *Handbook of Pragmatics* (pp. 607–632). Oxford: Blackwell Publishing.