Beyond ostension: Introducing the expressive principle of relevance

Constant Bonard

Institut Jean Nicod, ENS, Paris

ARTICLE INFO

ABSTRACT

In this paper, I am going to cast doubt on an idea that is shared, explicitly or implicitly, by most contemporary pragmatic theories: that the inferential interpretation procedure described by Grice, neo-Griceans, or post-Griceans applies only to the interpretation of ostensive stimuli. For this special issue, I will concentrate on the relevance theory (RT) version of this idea. I will proceed by putting forward a dilemma for RT and argue that the best way out of it is to accept that the relevance-theoretic comprehension procedure applies to certain non-ostensive stimuli, contrary to what is generally claimed within RT. In particular, I will argue that relevance theorists should accept that (ceteris paribus) non-ostensive emotional expressions in interactions guarantee a presumption of relevance such that they are interpreted through the relevance-theoretic comprehension procedure. This leads me to propose what I call 'the expressive principle of relevance'.

© 2021 Published by Elsevier B.V.

1. Introduction: a dilemma for relevance theory

The propositions (α), (β), and (γ) below are inconsistent. At least one of them must be rejected. However, as we will see, upholders of RT have good reasons to accept each. Therein lies the dilemma for RT.[1]

(α) If an interpretation procedure possesses the following properties (let us call them P), it is an instance of the relevance-theoretic comprehension procedure: a stimulus is interpreted by an audience through an inferential process guided by a presumption of relevance whereby the audience takes the information encoded in the stimulus and, thanks to its mindreading abilities, enriches and complements this information by following a path of least effort until the resulting interpretation meets its expectation of relevance.

(β) All and only stimuli perceived as ostensive are interpreted through the relevance-theoretic comprehension procedure.[2]

(γ) If stimuli perceived as ostensive are interpreted through a procedure with properties P, then some stimuli perceived as non-ostensive are also interpreted through it.

E-mail address: constant.bonard@gmail.com.

[1] Note that a version of this dilemma applies to all (neo-/post-)Gricean theories whose scope is restricted to ostensive stimuli (notably Horn, 1984; Levinson, 2000). The general structure of the dilemma is the following: (i) If a stimulus is interpreted through a procedure P, then it is an instance of (neo-/post-)Gricean procedure G. (ii) Only (and all) the stimuli that are perceived as ostensive are interpreted through (neo-/post-)Gricean procedure G. (iii) If stimuli perceived as ostensive are interpreted through procedure P, then some stimuli perceived as non-ostensive also are so interpreted.

[2] Although (β) already appears in the first edition of *Relevance* (1986), in later works Sperber and Wilson (2002; Wilson and Sperber, 2006) make a stronger claim: that the relevance-theoretic comprehension procedure is implemented by a Fodorian-like module which is activated exclusively by ostensive stimuli. (β) is more general: rejecting it implies rejecting the stronger claim. (β) thus makes the dilemma range broader.

I will detail what each means in the following three sections respectively (§2—4). Along the way, we will also see why they should be appealing to relevance theorists. Since (α) and (β) are already accepted in RT, making (γ) plausible will constitute the main goal of this paper. In §5, we will see that, as far as I can see, the best way out of the dilemma for relevance theorists is to reject (β),[3] despite the fact that it has been qualified as a central claim of RT (Sperber and Wilson, 1995, p. 50), and accept a modification instead, one that doesn't exclude non-ostensive emotional expressions from the RTCP. This will lead me to formulate what I call the 'expressive principle of relevance' (§5), a principle that complements the communicative principle of relevance.

2. The relevance-theoretic comprehension procedure

The relevance-theoretic comprehension procedure (RTCP for short) can be seen as a solution to the problem of code indetermination (a version of which is the problem of semantic underdetermination). Codes in this context are pre-established rules that associate information (in particular, messages) and stimuli (in particular, signals) and that must be mastered by communicators before the act of communication for it to take place successfully. Here is a very simple, famous code: light one lantern if the enemy comes by land, light two if it comes by sea. In this code, two pieces of information are respectively encoded into two types of stimuli. At another extreme of complexity, English grammar is a code that associates phonemes (stimuli) with semantic representations (information) according to pre-established rules (e.g. syntactic) that communicators using English must possess to achieve successful communication. Codes can also be non-conventional: vervet monkeys for instance make use of a genetically established code that pairs three kinds of calls (stimuli) with the presence of three kinds of predators (information).

The problem of code indetermination is that, in many instances of human communication, the relevant codes are insufficient to explain what information is transmitted. Consider the following (adapted from Wilson and Sperber, 2006, p. 615):

(1) — *Peter*: Did John pay back the money he owed you?
  — *Mary*: He forgot to go to the bank.

According to the relevant code — English syntax and compositional semantics — what Mary says in response to Peter's question is that John forgot to go to the bank. But, of course, we readily understand much more from her answer. We understand that John did not pay back the money he owed Mary, that Mary is not talking about a riverbank, but about a financial institution, and more. But how do we access these pieces of information since they are not encoded in the words that Mary uses? RT postulates that it is thanks to the RTCP.

To present the RTCP, let us go back to claim (α) of our dilemma. According to (α), if an interpretation procedure possesses the following eight properties, it is an instance of the RTCP (Sperber and Wilson, 2002, p. 18; Wilson and Sperber, 2006, p. 615ff):

(a) a stimulus is interpreted by an audience through an inferential process (usually unconscious, implicit, spontaneous, and fast)
(b) guided by a presumption of relevance whereby
(c) the audience takes the information encoded in the stimulus and
(d) thanks to its mindreading abilities
(e) enriches
(f) and complements this information
(g) by following a path of least effort
(h) until the resulting interpretation meets its expectation of relevance.

Let me make two remarks about these eight properties.[4] First, (a)-(h) are not to be conceived as sequential steps. For instance, to decode the signal (property (c)), we sometimes already need mindreading abilities (property (d)) and enrichment (property (d)), which shows that the temporal order may not follow the alphabetical order. Instead, the interpretative hypotheses may be made in parallel and adjusted to achieve a stable and sound inference in their final stage, perhaps by following 'mutual parallel adjustment' (Carston, 2002, pp. 143—146). Second remark: the claim in which we are interested is not that all eight properties are necessary for the RTCP. In some cases, some of the steps are not needed. Rather, (α) states that they are sufficient conditions for something to qualify as the RTCP.

Let us illustrate these eight properties with example (1) (from Wilson and Sperber, 2006, p. 615ff).

(a) The interpretation of Mary's utterance is not solely based on a pre-established code but goes through an inferential process, a form of abductive reasoning (unconscious, implicit, spontaneous, and fast) based on certain premises and the construction of (ad hoc) hypotheses.

---

[3] During the writing of this essay, Stravos Assimakopoulos sent me a draft (Assimakopoulos, 2021) which leads to a very similar conclusion, though it is not about emotional expressions.
[4] Thanks to the anonymous reviewers for remarking on these two points.

(b) One of the important premises in this inferential process is that Mary's utterance is relevant to the situation.
(c) We take the encoded (literal, semantic) meaning of her sentence as a cue to guess what she meant. As we saw, this is not sufficient.
(d) Another source of information is our mindreading capacities, i.e. our ability to ascribe mental states to others. Here, for instance, we ascribe to Mary the belief that it is common knowledge that banks can be financial institutions where one withdraws or sends money.
(e) As said above, the word 'bank' is ambiguous: it can mean river bank, investment bank, and retail bank. The word 'he' too is ambiguous as it may refer to any male individual. The process of understanding Mary's utterance goes through the enrichment of the bare code used by Mary in the sense that we must disambiguate the literal meaning of her sentence. According to the RTCP, we do so thanks to our mindreading abilities and our expectation that Mary's utterance is relevant.
(f) Once we understand what Mary literally means (Mary's explicature) notably by having resolved the ambiguities of the code, we further complement this information by making hypotheses as to why Mary would have said that. Because we assume her answer is relevant, we understand that no, John has not given her the money back.
(g) We may continue to extrapolate as to the possible implications of Mary's response, for instance by hypothesizing that either John does not have a great memory or that he was distracted. We may also infer that John won't repay Mary before he next goes to a bank. We do so by following a path of least effort in the sense that we test the hypotheses that appear as true or probably true that first come to our mind, i.e. in order of accessibility, and either accept or reject them.
(h) We stop inferring further conclusions once we are satisfied with an interpretation that makes Mary's utterance sufficiently relevant to us.

We should now have a clear enough idea of what the RTCP is. We may then ask: what are the reasons for a relevance theorist to reject (α)? Well, since there are, as far as I know, no definitions of the RTCP that conflict with (α), and since the conjunction of properties (a)-(h) seem to be sufficiently rich to identify RTCP and isolate it from other phenomena, I don't think there are any. In fact, even a person who disagrees with RT and believes that the way we interpret ostensive stimuli is wholly different from the process described by the RTCP has no reason to reject (α), because (α) just spells out a sufficient condition for a technical notion (namely the RTCP). It would be more reasonable for this person to accept (α) and further claim that the phenomenon it identifies does not correspond to the way we usually interpret ostensive stimuli.

3. RTCP and ostension

Ostensive stimuli are those stimuli that are produced with an informative and a communicative intention according to the following definitions:

"*Informative intention*: to make manifest or more manifest to the audience an array of propositions I.

Communicative intention: to make it mutually manifest to audience and communicator that the communicator has this informative intention." (Sperber and Wilson, 2015, p. 139).

This means that ostensive stimuli must make it mutually perceptible or inferable (Sperber and Wilson, 1995, p. 39) to the audience and communicator that they were produced with an informative intention. When we speak, this is usually obviously the case (unless we are speaking to ourselves, sleep-talking, rehearsing a discourse, etc.). We may also show our intention to communicate by many other means: if I gesture peculiarly while looking intensely into your eyes, you will probably guess that I want to inform you of something with this gesture. I could also draw something to you to make it manifest that I have an intention to inform you of something, or perform any other such action (on ostensive cues see e.g. Csibra, 2010).

Now, according to RT, there is an intimate link between ostension and RTCP. Here is a quote highlighting not only the intimateness but the uniqueness of the link between ostension and the inferential process postulated by RT[5]:

"The main thesis of this book is that an act of ostension carries a guarantee of relevance, and that this fact — which we will call *the principle of relevance* [NB: which is called the *Communicative Principle of Relevance* in the Postface and later writings] — makes manifest the intention behind the ostension. We believe that it is this principle of relevance that is needed to make the inferential model of communication explanatory." (Sperber and Wilson, 1995, p. 50).

The last sentence states that the inferential model of communication — which, in RT's version, includes the RTCP — cannot be explanatory without ostension. Why? Because without ostension, there is no presumption of *optimal relevance* — a claim that we will question below. A stimulus is optimally relevant if it is relevant enough to be worth the audience's processing effort (the interpretation) and if it is the most relevant one compatible with the communicator's abilities and preferences (Sperber and Wilson, 1995, p. 270). If a stimulus is ostensive, it carries a presumption of optimal relevance — that is what is

---

[5] To anticipate: I don't want to deny the existence of an intimate link between ostension and RTCP, but I am skeptical that this link is as *unique* as it is sometimes claimed to be. In particular, I find it plausible that the link between RTCP and emotional expressions is as intimate as that between RTCP and ostension (see §5).

stated by the Communicative Principle of Relevance (Sperber and Wilson, 1995, p. 270). If a stimulus is optimally relevant, then it is worth being processed through the RTCP. According to this reasoning, all stimuli perceived as ostensive are thus interpreted through RTCP.

But that is not all: we must infer from the last quote that it is not only *sufficient* that a stimulus carries a presumption of optimal relevance for it to be interpreted through RT's inferential process (and so the RTCP), but it is also *necessary*. Indeed, this is the only way to understand why Sperber and Wilson claim that the inferential model *cannot be explanatory* without the Communicative Principle of Relevance, i.e. without the presumption of relevance which comes with ostension.

Here is another quote that leads us to the same conclusion

> "There is thus a substantial difference between the frame of mind in which the individual may approach an ostensive stimulus directed at him and the frame of mind in which he approaches other phenomena. When attending to other phenomena, he may have hopes of relevance: if such hopes were totally unwarranted, there would be no point in attending to them at all. […] Even so, [with a non-ostensive stimulus] *there can be no general expectation of a steady and satisfactory level of relevance*. With an ostensive stimulus, however, the addressee can have not only hopes, but also fairly precise expectations of relevance." (Sperber and Wilson, 1995, p. 156 my emphasis).

The main point here is that, for any kinds of stimuli other than ostensive stimuli, we cannot expect a stable guarantee of sufficient relevance. In light of other writings (especially Sperber and Wilson, 2002; Wilson and Sperber, 2006), this means that we can rationally expect that ostensive stimuli are relevant enough to be worth the effort necessary to make the RTCP work, but we shouldn't expect any other kinds of stimuli to possess a level of relevance that is always high enough.[6]

We now have seen why relevance theorists may want to accept claim (β)[7]

(β) All and only stimuli perceived as ostensive are interpreted through the relevance-theoretic comprehension procedure.

Before we move on to the third claim, let me emphasize that, although RT has focused on ostensive stimuli and that (β) is a widespread assumption, the notion of relevance as it is defined by RT does not apply only to ostensive stimuli. Indeed, RT's cognitive principle of relevance applies to all cognitive inputs. For this reason, RT can be, and indeed has been, applied to non-ostensive stimuli (see notably some of the essays in Cave and Wilson, 2018; see also Dezecache et al., 2013). It is in this spirit that I will argue below that, given some assumptions about emotions, the cognitive principle of relevance allows us to derive not only the communicative principle of relevance, but another principle that applies to non-ostensive emotional expressions.

4. Non-ostensive stimuli and the RTCP

I want to suggest two ideas in this section. First, that the problem of code indetermination − understood in a broad sense which includes 'natural codes' (Wharton, 2003) − also arises with certain information exchanges through stimuli that are non-ostensive. Second, I will argue that if one accepts that the RTCP as it was described above is a satisfying solution to the problem of code indetermination for ostensive stimuli (§2), then one should also accept it as such for at least some non-ostensive stimuli, namely emotional expressions. We will see that, if that is correct, (γ) should be appealing to relevance theorists.

Let us concentrate on the following example:

(2) Jane really likes Maggie and this is weakly mutually manifest among them (Maggie suspects it, Jane suspects that Maggie suspects it, etc.). Jane and Maggie independently planned to go to a concert tonight. Maggie says "So, it turns out I have to replace Bob at the restaurant tonight …". Jane sighs. She didn't sigh with any informative or communicative intention and her sigh didn't carry any indication that she did (not particularly long or loud, etc.).

Let us make the plausible assumption that, with her sigh, Jane updates her and Maggie's mutual cognitive environment, i.e. the set of assumptions that are mutually manifest to them, with the following propositions:

- p: Jane is undergoing a negative affect (e.g. sadness or disappointment) about the fact that Maggie isn't coming.
- q: Both of them can now be pretty sure that Jane feels positively toward Maggie (in other words, it is now strongly mutually manifest to Maggie and Jane that Jane feels positively towards Maggie).

If Jane hadn't sighed, these pieces of information wouldn't have updated their cognitive environment, they would not have been inferable by Maggie and Jane, they would not have become mutually manifest as they have. How can we account for this matter of fact?

First of all, we cannot account for how Maggie and Jane's cognitive environment is updated merely by appealing to what sighs naturally mean (Grice, 1957). Contrary to the relation between the dark rings on a trunk and the age of the tree, or to that

---

[6] This seems to be one of the reasons that led Sperber and Wilson (Carston, 2002, pp. 143−146) to propose that the RTCP is implemented by a Fodorian-like module that is activated exclusively by ostensive stimuli, a claim which entails (β).

[7] Carston (2002, pp. 1−7) gives further support to (β) by arguing that it constitutes an answer to the skepticism of Davidson, Chomsky, and Fodor concerning pragmatics: if (β) is true, then pragmatics is not a 'theory of everything'. To anticipate, let me remark that this consequence also holds if (β*) is true (see §5 below). Thus, this argument does not favor (β) over (β*).

between certain spots and measles, there is not a factive relation between sighs and the propositions p and q. Even amended, probabilistic versions of natural meaning (for a review see Stegmanm (2015)) won't help here either since there is not a positive probabilistic correlation between sighs and p or q. It is not even the case that sighs naturally mean (statistically or not) that someone is undergoing a negative affect, as example (3) below will illustrate.

A more promising notion is that of signal (Hauser, 1996; Maynard Smith and Harper, 2003; Scott-Phillips, 2008; Skyrms, 2010). Signals — in contrast to non-communicative signs — are stimuli that have been designed by evolution to transmit information from a sender to a receiver. Wharton (2003, 2009) and Green (2007, 2019) successfully apply the notion of signal to many cases of emotional expression. Furthermore, as Wharton (2003) and Dezecache et al. (2013) argue, it is plausible that all emotional expressions are signals. If that is correct, we may explain how information about affective states is transmitted by appealing to how humans and other species are endowed with 'natural codes' for emotional expressions (Wharton, 2003, 2009). Natural codes are codes in the sense defined above — i.e. pairings of certain stimuli with certain information that are mastered by both the sender and the receiver in an act of communication. What is special is that the mastering of these codes is explained not by convention but by evolution — natural codes do not require conventionalized pairing between a symbol and some conceptual content.[8] This explains notably how non-human animals and infants can express and understand each other's emotions. See also Origgi and Sperber's (2000) analysis of bee communication through codes.

In our case however natural codes are not sufficient. This is because, if we look at the pre-established pairings between sighs and the pieces of information they carry — sighs' code —, these pairings won't be sufficient for Maggie to infer that p and q. Sighs, just like most — and perhaps all — types of emotional expressions, can express many different mental states. It has been shown that sighs can signal pain, panic, relaxation, relief, sadness, stress, and the will to give up (Teigen, 2008; Vlemincx et al., 2009). Furthermore, sighs' code doesn't tell us what the sighs are about. If we just look at what is encoded in the sigh, we won't find what we are looking for — namely p and q.

To reinforce this last conclusion, consider the following variation on the last scenario:

(3) Exactly as (2) except that Jane *doesn't* like Maggie at all.

In this case, it would be reasonable to consider that Jane updates her and Maggie's mutual cognitive environment with the following pieces of information:

- r: Jane undergoes a *positive* affect (e.g. relief) about the fact that Maggie isn't coming.
- s: Both of them can now be pretty sure that Jane feels *negatively* toward Maggie.

Comparing (2) and (3) makes it obvious that the information encoded into sighs is very ambiguous. What a sigh encodes is that the person is either relieved, or disappointed, or tired, or stressed, or in another affective state with which sighs are correlated. In other words, we are once again faced with a problem of code indetermination, just as we were with Mary's utterance in (1). Both are cases of code indetermination in the sense that the information carried by the stimuli themselves is too ambiguous to explain what information is in fact transmitted. The information that these stimuli encode (whether it be a string of words or a sigh) is too vague, too indeterminate to account for what we understand from them, for how they update our mutual cognitive environments.[9]

Now, for RT, the big difference between, on the one hand, (2) and (3) and, on the other hand, (1) is that Jane's sighs are *not* ostensive stimuli. As the examples specify, the sighs were not produced with a communicative and an informative intention.

Spontaneous emotional expressions may well be used as ostensive stimuli (Wharton, 2003, 2009), but they very often are not (Dezecache et al., 2013). Examples (2) and (3) belong to the latter case. (2) and (3) can advantageously be conceived as cases of *non-ostensive showing*. Despite being non-ostensive, Jane's sighs indeed show her emotions in the sense that they reveal or display them and, plausibly, have the function to do so (for a detailed discussion of showing in this sense, see Green, 2007, Chapters 1—3). Let me say more about this.

Relevance theorists have insightfully argued that, contrary to what Grice (1957) suggested, there is a continuum of cases between ostensive showing and meaning non-naturally (meaningNN) without clear boundaries (Sperber and Wilson, 2015; Wharton, 2003, 2009; Wilson and Wharton, 2006). However, if this showing-meaningNN continuum only applies to ostensive stimuli, then the sighs in (2) and (3) do not belong to it.[10] Indeed, because Jane sighs without any communicative and informative intention, (2) and (3) are to be contrasted with cases where the sender ostensively shows an emotional or physiological response, such as when a child deliberately and openly shows to her mother how pale she is or when Sam ostensively shows that he is shivering to communicate that he is freezing (Grice, 1989, p. 218; Sperber and Wilson, 2015; Wharton, 2003, 2009,

---

[8] It is of course an empirical question whether all emotional expressions possess a communicative function and so whether we should talk about, for instance, sighs as signals that possess a natural code or as non-communicative signs that possess a natural statistical meaning. My argument can be made in either case but, in order to avoid lengthy disjunctive formulations, it will be more convenient to follow Wharton as well as Dezecache et al. and assume that sighs are signals. Nevertheless, we can keep in mind that a phrase such as 'Sighs encode x' could be rephrased as 'Sighs naturally mean x (in the statistical sense)' and that the 'problem of code indetermination' could be rephrased as 'the problem of indetermination in what stimuli naturally mean in the statistical sense'.

[9] Observe that this notion of indetermination does not depend on the fact that the communicator had informative or communicative intentions.

[10] RT has been applied to non-ostensive emotional stimuli, notably by Dezecache et al. (2013) and Sellevold (2018). However, the latter works are about a different problem — that of emotional vigilance. They do not discuss the relation between non-ostensive emotional expressions and the RTCP or if these stimuli have a presumption of relevance.

Chapter 2; Wilson and Wharton, 2006). Jane's sighs are also to be contrasted with cases where it is ambiguous whether one intends to reveal something with one's physiological or emotional reaction, such as the blushing of Lily Bart in Edith Wharton's novel *The Art of Mirth* (1905). We may hypothesize that Lily, at least sometimes, use her blushing intentionally as a communicative tool (Sellevold, 2018). (2) and (3) are not to be conceived as such: in these examples, it is clear that Jane sighed non-ostensively. The sighs are not either non-ostensive modifications of ostensive stimuli (such as non-ostensive prosodic modifications of verbal utterances, see Wilson and Wharton (2006)). They are fully non-ostensive stimuli. Furthermore, observe that we don't need to perceive or conceive of the sighs in (2) and (3) as ostensive stimuli to infer p and q or r and s.

In sum, the problem of code indetermination also arises with certain information exchanges through stimuli that are non-ostensive and perceived as non-ostensive. This was the first point that I wanted to make in this section.

It is easy to multiply the examples where spontaneous emotional expressions convey more than the information they encode. Think, for instance, about all the different affective states with which low pitched, soft laughter can be associated — a review of the empirical literature reveals that such laughter correlates with amusement, contempt, fear, incredulity, joy, sadness, Schadenfreude, social anxiety, urge to affiliate, urge to aggress, and ticklishness (Bonard, 2021, sec. 1.5). Even Ekman's supposedly universal facial expressions of emotion are interpreted to express different emotions in different contexts (see e.g. Barrett et al., 2011).

Furthermore, even when we can infer what kind of emotion a stimulus expresses, what the stimulus itself reveals usually does not allow us to infer what the emotion is about. Nevertheless, we normally automatically get what emotional states are expressed and what they are about. More generally, there appear to be no unambiguous emotional expressions, although we usually understand very precise information from them.[11]

In all these cases, I surmise, we complete the information encoded in the emotional expression through mindreading, inferential processes. In fact, I believe that the inferential process used to infer p—s is the same as the one used to infer the meaning of ostensive stimuli. In particular, as far as RT is concerned, I would hold the following, which is the third claim of our dilemma:

(γ) If stimuli perceived as ostensive are interpreted through a procedure with properties P, then some stimuli perceived as non-ostensive also are so interpreted.

You will remember that the properties P are the ones which were discussed in §2:

(a) the stimulus is interpreted by an audience through an inferential process (usually unconscious, implicit, spontaneous, and fast)
(b) guided by a presumption of relevance whereby
(c) the audience takes the information encoded in the stimulus and
(d) thanks to its mindreading abilities
(e) enriches
(f) and complements this information
(g) by following a path of least effort
(h) until the resulting interpretation meets its expectation of relevance.

Let me take (2) as an illustration of how the inferential process used to interpret certain non-ostensive stimuli possesses these eight properties. You will notice that this illustration parallels quite strictly the one given for (1) above (§2).

(a) The interpretation of Jane's sigh is not solely based on a pre-established (natural) code but goes through an inferential process, a form of abductive reasoning (unconscious, implicit, spontaneous, and fast) based on certain premises and the construction of (ad hoc) hypotheses.
(b) One of the important premises in this inferential process is that Jane's sigh is relevant to the conversation. First, the sigh does not come out of the blue but is an affective reaction to Maggie's utterance (this is comparable to the fact that we take Mary's utterance in (1) to be a *response* to Peter's question and so as relevant to it, although, importantly, the ostensive element is absent from Jane's sigh). Second, as we will see below, because it is an affective response, we can take it to indicate that Maggie's utterance is a highly relevant stimulus for Jane.
(c) We take the encoded meaning of her sigh as another premise to guess what information she is thereby transmitting. The relevant code tells us that her sigh indicates that she underwent (or realistically faked a spontaneous expression of[12]) relief, sadness, relaxation, an urge to abandon, or another affective response to what Maggie said (see above). The

---

[11] Observe that what I claim here is compatible with the existence of a 'core meaning' for emotional expressions, parallel to lexically encoded meaning, i.e. some information that is encoded in or naturally meant by all tokens of a given type of emotional expression (thanks to an anonymous reviewer for this remark). The expression 'the information encoded in emotional expressions' is meant to refer to this core meaning (see footnote 9).

[12] If Jane realistically faked a spontaneous expression her emotion, she had an informative intention, contrary to case (2) and (3). However, if she did not have a communicative intention because she intended to hide her informative intention. This case is different to an ostensive mimicry (e.g. an ostensively exaggerated sigh) where one has an communicative intention.

fact that she underwent an affective response also carries the encoded information that she cared about Maggie's utterance (more on this below). This decoding is very ambiguous though − someone who is engaged in the conversation will want to know *which* affective state Jane undergoes and *what* it is about.

(d) Another source of information is our mindreading capacities, i.e. our ability to ascribe mental states to others. We ascribe to Jane a liking for Maggie and, relatedly, suppose that an important goal for her is to spend more time with Maggie, especially outside work. We also attribute to her the inkling that Maggie knows she likes her.

(e) Because sighs are ambiguous, the process of understanding Jane's sigh goes through the *enrichment* of the bare code in the sense that we disambiguate what is encoded in the sigh. Thanks to our expectation that Jane's sigh is a relevant reaction to Maggie's utterance (as opposed to an unrelated reaction such as a sneeze) and thanks to our mindreading abilities, we infer that Maggie is probably disappointed and/or saddened.

(f) Once we have disambiguated in what psychological state Jane is, we *complement* this information by making further hypotheses, for instance concerning what she is sad or disappointed about. This leads us to infer that both of them can now be pretty sure that Jane feels positively toward Maggie.

(g) We may continue to extrapolate as to the possible implications of Jane's sigh, for instance by hypothesizing that if Maggie proposed that they go to another gig next week, she would be very glad. We may also hypothesize that, if she learns that Maggie does not feel the same way toward her, Jane may be embarrassed by her sigh (because it showed her feelings). We evaluate the hypotheses that appear as true or probably true by following a path of least effort in the sense that we either accept or reject the ones that first come to our mind in order of accessibility (or manifestness).

(h) We stop inferring further conclusions once we are satisfied with an interpretation that makes Jane's sigh sufficiently relevant to us. If, for instance, Maggie does not care a lot about what Jane thinks, she may not find it worth it to make the effort to draw further conclusions than p and q. If, on the other hand, she does care about her, she may make many more hypotheses than the ones discussed, because the cognitive cost of doing this inferential work would be compensated by what she considers to be highly valuable information.

It seems to me that, for those who find it plausible that we interpret ostensive stimuli as in (1) through an inferential process with properties (a)–(h), it should be just as plausible that we interpret Jane's sigh through an inferential process with the same properties, even though Jane's sigh is non-ostensive. This is, in substance, claim (γ) of our dilemma. I believe that the considerations in this section make this claim attractive.

5. Facing the dilemma: The expressive principle of relevance

Here is, once again, the dilemma with which, I have argued, relevance theorists are faced:

(α) If an interpretation procedure possesses the properties P (see §2), it is an instance of the relevance-theoretic comprehension procedure.

(β) All and only stimuli perceived as ostensive are interpreted through the relevance-theoretic comprehension procedure.

(γ) If stimuli perceived as ostensive are interpreted through a procedure with properties P, then some stimuli perceived as non-ostensive also are so interpreted.

I would reject (β). Instead, I think it should be replaced with the following:

(β*) All stimuli perceived as ostensive are interpreted by the audience through the relevance-theoretic comprehension procedure. But other kinds of stimuli are so interpreted as well, for instance, stimuli perceived as non-ostensive emotional expressions in an interaction.

There are three reasons for this proposal. One is that (β*) is compatible with (α) and (γ), both of which I find more plausible than (β). The second is that the arguments that were presented in §3 for accepting (β) only make the first sentence of (β*) convincing. The third is that emotional stimuli, just like ostensive stimuli, guarantee a presumption of relevance. Let me say a few words about the second reason before I turn to the third.

We saw that ostensive stimuli always convey a presumption of optimal relevance and that, consequently, all stimuli perceived as ostensive carry a presumption that they are worth the audience's effort of processing them with the RTCP. But this only leads to the claim that it is *sufficient* that a stimulus is perceived as ostensive for it to be processed by the RTCP, which is what the first sentence in (β*) states. The reasoning leading to (β), i.e. that ostension is sufficient *and* necessary for the RTCP, was that there are no other kinds of stimuli that are stable guarantees of sufficient relevance; only ostensive stimuli always carry a presumption of being worth the effort of going through the RTCP. However, no positive arguments were given for this stronger claim. Aren't there other kinds of stimuli that carry such a presumption of relevance? As we will see, emotional stimuli are a candidate for this status.

A worry that some readers may have at this point is that (β*) opens the door of pragmatics to any old kind of non-ostensive stimuli, so that pragmatics would turn into a theory of everything, an objection raised by Davidson, Chomsky, and Fodor (Carston, 2002, p. 1). This worry is not warranted because there is something special about emotional stimuli that makes them

comparable to ostensive stimuli, as we will see. This can be formulated with the following principle, which is complementary to the communicative principle of relevance:

Expressive principle of relevance

Any emotional expression in an interaction carries a presumption of (quasi-)optimal[13] relevance, i.e. it is relevant enough for it to be worth the interactant's effort to process it through pragmatic mechanisms, in particular the effort necessary for the RTCP.[14]

I will now sketch the rationale for this principle. This can only be a tentative argument and more work would be required to sustain my claims but a detailed argumentation and an exploration of its consequences would require a paper of its own (Wharton et al., 2021 give arguments that support some of the following). Nevertheless, I hope that the following will suffice in indicating how the expressive principle of relevance may be derived from what we know about emotions together with the cognitive principle of relevance.

It has become a consensus in affective sciences that emotions are reactions to stimuli perceived as highly pertinent[15] to the person undergoing the emotion, in the sense that the stimuli are perceived as obstructive or conducive to important goals or concerns, so that the stimuli in question demands an emotional response (Lange et al., 2020; Sander, 2013; Scherer and Moors, 2019). In fear, for instance, the fear-inducing stimulus is perceived (unconsciously) as a threat to one's safety or wellbeing and so as demanding a reaction such as escape or preventive attack, which itself is prepared by the physiological modifications accompanying fear (increased blood flow, increased breathing, tensed muscles, etc.).

Because (ceteris paribus[16]) all emotional episodes are reactions to stimuli appraised as highly pertinent to her goals, when someone reacts emotionally in an interaction, this indicates that something in the context of the interaction is perceived as highly pertinent. What is highly pertinent to the person with whom we interact may well be pertinent to us as well, either because it may directly concern us (e.g. a threat to her may be a threat to us) or because it will have important consequences on the course of the interaction, notably by strongly modifying what is manifest to our interactant's mind, since the objects of emotions are (ceteris paribus) highly manifest to the person undergoing them.

Since, according to the cognitive principle of relevance (Sperber and Wilson, 1995, p. 260), human cognition tends to be geared towards the maximization of relevance, i.e. to the achievement of as many cognitive effects as possible for as little processing effort as possible, if one detects in the person with whom one is interacting an emotional reaction, because emotions are reactions to stimuli appraised as highly pertinent, this should trigger an attempt of the audience to hypothesize what in the context of the interaction is so highly pertinent to the interactant. Not putting in the effort to interpret what emotional state is expressed and what in the context of the interaction triggered this emotional state could have important negative consequences, because the audience may be left unaware of something very important for their interaction and beyond. For instance, in (2) and (3), if Maggie does not try to figure out what emotional state Jane is in and what it is about, she would miss a precious opportunity to gain some very important information about their present interaction and their personal relation in general. In virtue of the nature of emotions, the cognitive effects to be gained from the interpretation of the emotional expression of the person with whom one is interacting prima facie make it worth the effort necessary for the interpretation.

Now, since all emotional expressions are ambiguous (in the sense that there is no code that pairs them with an emotional state and what this state is about) and that, for reasons given in §4, the disambiguation of emotional expressions should be done by an interpretation that possesses the properties (a)-(h), all emotional expressions in interactions are relevant enough to be worth the effort necessary for the RTCP (I thus deny, as per (β*), that only ostensive stimuli trigger the RTCP).

Let me now say why I use the expression 'presumption of (quasi-)optimal relevance' instead of 'presumption of optimal relevance'. The official definition of the latter is the following:

(a) "The ostensive stimulus is relevant enough for it to be worth the addressee's effort to process it.
(b) The ostensive stimulus is the most relevant one compatible with the communicator's abilities and preferences." (Sperber and Wilson, 1995, p. 270, p. 270)

I have argued that an emotional expression in an interaction, even when it is not ostensive, is relevant enough for it to be worth the addressee's effort to process it. I have not however given an argument for the claim that it is the most relevant

---

[13] I will indicate below why I use '(quasi-)optimal'.

[14] In case other accounts of the inferential interpretations turn out to be correct, the mention to the RTCP here should be replaced by whatever inferential process is supposed to play the role that the RTCP plays within RT today.

[15] Normally, the word 'relevant' rather than 'pertinent' is used in this context. I use 'pertinent' to avoid presupposing that it is the same as RT's relevance. However, see Wharton et al. (2021) for the idea that 'relevance' in RT and in affective sciences refer to the same property.

[16] On the status of 'ceteris paribus', see Fodor (1987, Chapter 1).

stimulus compatible with the communicator's abilities and preferences and leave this question open.[17] That is why I prefer to use the expression 'presumption of (quasi-)optimal relevance' to qualify the fact that all emotional expressions in interactions are relevant enough to be worth the effort necessary for the RTCP.

I have explained why, given the nature of emotions and the cognitive principle of relevance, it makes sense to hold the expressive principle of relevance. The argument is less straightforward than the one leading from the cognitive principle to the communicative principle of relevance, but it nevertheless appears to be sound.

Besides this theoretical argument, empirical evidence may be found in support of − or against − this principle, since it leads to testable hypotheses. For instance, it predicts that (ceteris paribus) when two people are engaged in a conversation, if one of them reacts with a detected non-ostensive emotional expression, the interlocutor will infer from this emotional expression more than what is encoded in it. It further predicts that the interlocutor would do so thanks to mental mechanisms defining the RTCP.

6. Other ways out of the dilemma?

In the last section, I have proposed my favored solution to the dilemma and argued that the expressive principle of relevance would explain why we should accept (β*) instead of (β). Before I finish, let me briefly mention other options for relevance theorists that I find less satisfying and respond to two extra objections.

One may reject (α) instead of (β) by arguing that the properties (a)-(h) would not be sufficient to identify the RTCP and that it would need to be characterized in more detail, notably by referring to explicit and implicit content, intended contextual assumptions and implications, the Communicative Principle of Relevance, and/or other features (for these notions, see Sperber and Wilson, 1995 Postface; Wilson and Sperber, 2006). Once identified with further properties, the RTCP won't apply anymore to Jane's sigh. One can thus maintain (β) and (γ) by arguing that, even if the argument in section §4 is correct, the way we interpret Jane's sigh is not through the RTCP, but merely through a *similar* inferential process.

The problem with this option is that it postulates two largely redundant mental processes used for interpretation (for a similar point, see Assimakopoulos, 2021, sec. 3). Indeed, it would mean that two distinct mental processes possess properties (a)−(h), but that one of them − the "true" RTCP − has further properties. However, it seems much more plausible that there is only one kind of inferential process and that, with certain stimuli, specialized mechanisms are activated in addition (e.g. mechanisms specialized for verbal interpretation). This is especially striking if we accept that the RTCP is implemented by an innately-determined module (Sperber and Wilson, 2002). Why would evolution, which usually is parsimonious, have selected two different brain mechanisms to do roughly the same job? In other words, the interpretation process described in §2 seems to be largely the same as the one described in §4, just as the illustration of points (a)−(h) in the two cases seems to show. But then, why wouldn't relevance theorists accept (α) and call both of them RTCP?

Another option that I find unsatisfying is to reject (γ). A relevance theorist may want to do so by arguing that my descriptions of how properties (a)−(h) apply to the interpretation of Jane's sigh are implausible. However, I don't see why the application of (a)−(h) to Jane's sigh would be any less plausible than to Mary's utterance in (1). Another way to attack (γ) would be to argue that Jane's sigh is perceived as ostensive. The problem with this claim is that it is widely agreed, including by relevance theorists (Dezecache et al., 2013), that spontaneous emotional expressions can be produced without the intentions defining ostensive stimuli.[18] Why wouldn't Jane's sigh be perceived as non-ostensive?

In any case, rejecting (γ) leads to a problem: if it is not through the RTCP that we interpret non-ostensive stimuli whose codes underdetermine their meaning, then how do we interpret them? And what is faulty about the argument in §4? These are, I believe, important challenges for the relevance theorists who wouldn't accept (γ). I consider that the burden of the argument now is on their shoulders.

Before I turn to the conclusion, let me address two objections formulated by anonymous reviewers (for which I thank them very much). The first one is that, although the RTCP has as its proper domain ostensive behavior, it will inevitably respond to other types of stimuli that are not its proper domain as if they were because they somehow resemble the latter. We can think for instance about the thought experiment by Knapp and Michael (1982) where waves and other natural phenomena have somehow created forms on the beach that look like a poem by Woodsworth: this is not an ostensive stimulus but it will naturally trigger the RTCP.

Of course, this can sometimes happen but, if I am correct in my prediction that (ceteris paribus) all non-ostensive spontaneous emotional reactions in an interaction trigger the RTCP, then it is not plausible that they always do so because of an accidental resemblance with non-ostensive behavior, because non-ostensive emotional reactions can be clearly different from ostensive stimuli (e.g. blushing). As such, the comparison with Knapp and Michael's poem on the beach does

---

[17] Here is a tentative argument for this claim: suppose first that the communicator cannot prevent herself from expressing her emotion, that it is irrepressible. In this case, the communicator's abilities prevent her from displaying any other stimuli. Suppose, now that the communicator could have prevented herself from displaying her emotion. If she considered that this stimulus was not the most relevant one compatible with her preferences, (ceteris paribus) she would have refrained herself from producing it, since she had the ability to do so. In both cases, the stimulus appears to be the most relevant one given her abilities and/or her preferences. Here is another small argument: some may consider that the RTCP is triggered only by stimuli that carry an assumption of optimal relevance (this was suggested by the remark of an anonymous reviewer although, to the best of my knowledge, this is not stated explicitly in the RT literature). If that is correct, if (α) is correct (§2), and if the interpretation of non-ostensive emotional expressions in interaction possess properties (a)−(h) (§4), then it follows that non-ostensive emotional expressions in interactions carry a presumption of optimal relevance.

[18] As discussed above, they sometimes are shown ostensively, but it is not the case with Jane.

not hold because the latter clearly resembles an ostensive stimulus. Non-ostensive emotional reactions can appear as clearly non-ostensive and, I have argued, nevertheless be stable triggers of the RTCP, so that an accidental trigger due to a putative resemblance to ostensive stimuli seems an unsatisfying explanation of the phenomenon.

The second objection is that, if we were to follow my proposal, we may run a risk of having to assume a very large number of ad-hoc principles of relevance for any kind of non-ostensive stimulus that is by default treated as having potentially high relevance, e.g. principles of relevance applying to bright-colored clothes, that someone comes early for their meeting, that someone doesn't make eye-contact, etc. For instance, what would prevent us from extending my argument to defend that we need a principle stating: "Any loud noise in the environment carries a presumption of relevance that is strong enough for the participants in the conversation to put in the effort necessary for an inferential interpretation."

I respond to this interesting objection by highlighting that, if we follow my proposal, we do not run the risk of devising a multitude of ad-hoc principles because emotional expressions, like ostension, have a special relation to relevance and communication, a relation that, prima facie, is rare and potentially unique to these two types of stimuli. The special relation that relevance and communication possess with ostension is extensively discussed in RT. The one they have with emotional expressions was sketched in the last section. To sum this up in a nutshell, emotional expressions carry a stable presumption of relevance — i.e. we can reliably expect that its positive cognitive effects are worth the processing efforts — in virtue of the fact that emotions are psychological states that involve an appraisal of the situation as being highly conducive or obstructive to one's goals, and that, in the context of an interaction with someone displaying an emotion, interpreting what the emotion is and what it is about is always worth the effort. It is not the case that bright-colored clothes (besides the fact that clothes arguably are ostensive stimuli), early arrivals to a meeting, or loud noises carry a stable guarantee of relevance. For instance, there are many contexts where loud noises are not relevant enough for us to spend the effort of trying to interpret them, because, say, they are frequent and signal something benign in this context (e.g. in factories). Furthermore, contrary to emotional expressions, such stimuli generally do not require an interpretation process with properties (a)-(h). In sum, stimuli that are often attention-grabbing need not possess a stable presumption of relevance that is high enough to be worth an interpretation, not to mention an interpretation through the RTCP. By contrast, I argue that emotional expressions in inter-action do possess it. As such, they play a role comparable to ostensive stimuli in communication (and they do so not because they resemble ostensive stimuli, as noted above). Maybe other kinds of stimuli also qualify, but I cannot think of any.

7. Conclusion

Although some of the claims defended here are in contradiction with tenets of RT, all in all, I take it to be good news for RT because, if I am correct, it would also mean that it could expand its tools — in particular the RTCP — to further territories: the interpretation of non-ostensive emotional expressions. This prospect is especially interesting for RT since, to the best of my knowledge, these territories have for the most part remained unexplored (though see e.g. Dezecache et al., 2013; Sellevold, 2018) and since the affective sciences of the past decades have provided a wealth of data and theories waiting to be integrated to pragmatics.

Declaration of competing interest

The author has no conflicts of interest to declare.

Acknowledgements

References

Assimakopoulos, S., 2021. Beyond meaningNN and ostension: pragmatic inference in the wild. In: Wharton, T., de Saussure, L., Ifantidou, E. (Eds.), Beyond Meaning. John Benjamins.
Barrett, L.F., Mesquita, B., Gendron, M., 2011. Context in emotion perception. Curr. Dir. Psychol. Sci. 20 (5), 286—290.
Bonard, Constant, 2021. Meaning and emotion. Doctoral dissertation. University of Geneva and University of Antwerp. https://doi.org/10.13097/archive-ouverte/unige:150524.
Carston, R., 2002. Thoughts and Utterances: the Pragmatics of Explicit Communication. Blackwell.
Cave, T., Wilson, D. (Eds.), 2018. Relevance Theory and Literary Interpretation. Oxford University Press.
Csibra, G., 2010. Recognizing communicative intentions in infancy. Mind Lang. 25 (2), 141—168.
Dezecache, G., Mercier, H., Scott-Phillips, T.C., 2013. An evolutionary approach to emotional communication. J. Pragmat. 59, 221—233.
Fodor, J., 1987. Psychosemantics: the Problem of Meaning in the Philosophy of Mind. MIT Press pp. xiii, 171.
Green, M., 2007. Self-Expression. Oxford University Press.
Green, M., 2019. Organic meaning: an approach to communication with minimal appeal to minds. In: Capone, A., Carapezza, M., Lo Piparo, F. (Eds.), Further Advances in Pragmatics and Philosophy: Part 2 Theories and Applications. Springer International Publishing, pp. 211—228. https://doi.org/10.1007/978-3-030-00973-1_12.
Grice, H.P., 1957. Meaning. Phil. Rev. 66 (3), 377—388.
Grice, H.P., 1989. Studies in the Way of Words. Harvard University Press.
Hauser, M.D., 1996. The Evolution of Communication. MIT Press.

Horn, L., 1984. Toward a New Taxonomy for Pragmatic Inference: Q-Based and R-Based Implicature. *Meaning, Form, and Use in Context: Linguistic Applications*, p. 11.

Knapp, S., Michaels, W.B., 1982. Against theory. Crit. Inq. 8 (4), 723—742.

Lange, J., Dalege, J., Borsboom, D., van Kleef, G.A., Fischer, A.H., 2020. Toward an integrative psychometric model of emotions. Perspect. Psychol. Sci. 15 (2), 444—468. https://doi.org/10.1177/1745691619895057.

Levinson, S.C., 2000. Presumptive Meanings: the Theory of Generalized Conversational Implicature. MIT Press.

Maynard Smith, J., Harper, D., 2003. Animal Signals. Oxford University Press.

Origgi, G., Sperber, D., 2000. Evolution, communication and the proper function of language. In: Carruthers, P., Chamberlain, A. (Eds.), Evolution and the Human Mind: Language, Modularity and Social Cognition. Cambridge University Press, pp. 140—169.

Sander, D., 2013. Models of emotion. In: Armony, J., Vuilleumier, P. (Eds.), The Cambridge Handbook of Human Affective Neuroscience. Cambridge University Press, pp. 5—54. https://doi.org/10.1017/CBO9780511843716.003.

Scherer, K.R., Moors, A., 2019. The emotion process: event appraisal and component differentiation. Annu. Rev. Psychol. 70 (1), 719—745. https://doi.org/10.1146/annurev-psych-122216-011854.

Scott-Phillips, T., 2008. Defining biological communication. J. Evol. Biol. 21 (2), 387—395.

Sellevold, K., 2018. On the borders of the ostensive: blushing in Edith Wharton's the House of Mirth. In: Cave, T., Wilson, D. (Eds.), Reading beyond the Code: Literature & Relevance Theory. Oxford University Press, pp. 91—108.

Skyrms, B., 2010. Signals: Evolution, Learning, and Information. Oxford University Press.

Sperber, D., Wilson, D., 1995. In: Relevance: Communication and Cognition, second ed. Blackwell.

Sperber, D., Wilson, D., 2002. Pragmatics, modularity and mind-reading. Mind Lang. 17 (1—2), 3—23.

Sperber, D., Wilson, D., 2015. Beyond speaker's meaning. Croat. J. Philos. 15 (44), 117—149.

Stegmann, U.E., 2015. Prospects for probabilistic theories of natural information. Erkenntnis 80 (4), 869—893.

Teigen, K.H., 2008. Is a sigh "just a sigh"? Sighs as emotional signals and responses to a difficult task. Scand. J. Psychol. 49 (1), 49—57.

Vlemincx, E., Van Diest, I., De Peuter, S., Bresseleers, J., Bogaerts, K., Fannes, S., Li, W., Van Den Bergh, O., 2009. Why do you sigh? Sigh rate during induced stress and relief. Psychophysiology 46 (5), 1005—1013.

Wharton, T., 2003. Natural pragmatics and natural codes. Mind Lang. 18 (5), 447—477.

Wharton, T., 2009. Pragmatics and Non-verbal Communication. Cambridge University Press.

Wharton, Tim, Bonard, Constant, Dukes, Daniel, Sander, David, Oswald, Steve, 2021. Relevance and emotion. J. Pragmat. In this issue.

Wilson, D., Sperber, D., 2006. Relevance theory. In: Horn, L. (Ed.), The Handbook of Pragmatics. Blackwell.

Wilson, D., Wharton, T., 2006. Relevance and prosody. J. Pragmat. 38 (10), 1559—1579.

Constant Bonard is a postdoctoral researcher in philosophy at the Institut Jean Nicod (École Normale Supérieure, Paris), grantee of the Swiss National Science Foundation. He works on emotion, on communication, and especially on the intersection of these two topics. Constant Bonard also conducts research in music psychology and in aesthetics.