

Pragmatics, cognitive heuristics and the straw man fallacy

Steve Oswald, University of Fribourg & University of Neuchâtel
Marcin Lewiński, Universidade Nova de Lisboa

1. Introduction

In the literature on argumentation, the straw man fallacy denotes the misrepresentation of someone's position in order to easily refute that position (see Aikin & Casey 2011, Lewiński 2011, Bizer, Kozak, & Holterman 2009, van Eemeren & Grootendorst 1992: Ch. 11, van Laar 2008, Talisse & Aikin 2006, Walton 1996, Walton & Macagno 2010). An argumentative move that involves a straw man can thus be characterised by its two constitutive elements (Lewiński 2011): the function of refuting an opponent's position and the form consisting in various methods of misrepresentation of the original position (misquotation, selective quotation, taking out of context, attacking a fictitious opponent, etc.). The straw man fallacy has drawn considerable scholarly attention, not least because of its prevalence in public discourse (see esp. Aikin & Casey 2011 and Talisse & Aikin 2006).

What is a remarkable but often overlooked feature of the straw man is that, contrary to most other common fallacies such as *ad hominem* or *ad baculum*, it is a *meta*-discursive fallacy. It “operates” on someone else's discourse that serves as material for linguistic manoeuvring.¹ It therefore does not neatly fit into the Aristotelian categories of verbal (“dependent on language”) and material (“not dependent on language”) fallacies (see

1 As mentioned above, this includes purported discourse attributed to some fictitious or unidentified/unidentifiable “author”: frequent examples include clauses such as “People out there may say P”, “It is well-known that P”, etc. In this variant, the question of who exactly these “people” are becomes crucial.

Hamblin 1970).² A straw man attack involves a certain unreasonable, sometimes manipulative, re-interpretation of another position. Hence, it plays on the interpretative resources of arguers and its identification as such requires careful meta-linguistic (or meta-discursive) attention.

In our previous work (Lewiński & Oswald 2013) we argued that to fully grasp both the pragmatic and the argumentative mechanisms behind the straw man fallacy, one has to answer two questions: (i) *when* can we justifiably say that the straw man occurred? and (ii) *how* can we explain its remaining covert and thus its persuasive or even manipulative potential? We posited that a satisfactory answer requires a theoretical combination of two branches of pragmatic inquiry: a *normative* pragmatic theory of argumentation, such as Pragma-Dialectics (henceforth PD), and a *cognitive* pragmatic model of meaning (and in particular of the cognitive mechanisms governing interpretation), such as the one offered by Relevance Theory (henceforth RT). Our rationale for such combination was that the normative pragmatic component of PD provides a framework for answering question (i), while RT's empirically adequate cognitive pragmatic account is capable of addressing the problems involved in question (ii). We found these two theories complementary in their goals and methods for the *analysis* of argumentative discourse. We argued that this is clear in both their pragmatic take on meaning in discourse and their heavy reliance on the notion of context. Within this hybrid framework, we stipulated contextual, pragma-dialectical conditions for the identification of the alleged straw man attacks and provided a basic cognitive account of the manipulative effectiveness of straw man abuses (see also section 3.2).

This contribution builds on our previous results. Its goal is twofold: first, we aim to refine the grounds on which we claim both theories can be combined in argumentative, and more specifically rhetorical, analysis and, second, extend our cognitive account, so that question (ii) above becomes more prominent than question (i) which was primarily addressed in our previous research. We will thus carefully explore the pragmatic relation between the normative violation that makes a criticism fallacious and the

2 It may be taken as a conversational sub-category of *ignoratio elenchi* of sorts, and thus as a fallacy not dependent on language. However, it clearly depends on linguistic means of misrepresentation, which would suggest an opposite classification. Note that neither the straw man nor "ad" fallacies were discussed by Aristotle (see Hamblin 1970).

cognitive “mistake” that makes an unreasonable straw man acceptable to reasonable hearers. More specifically, we will investigate the cognitive mechanisms responsible for an addressee’s failure to identify the straw man as an unreasonable, fallacious, critique.

To address these questions, we will proceed as follows. First, in section 2 we will critically review the most relevant attempts within Argumentation Theory to account for the treacherousness of fallacies in terms of simplified cognitive heuristics and elaborate on the cognitive underpinnings of argument processing following recent findings in the cognitive study of epistemic aspects of communication (see Sperber *et al.* 2010). Second, in section 3 we will highlight some issues raised by Walton’s account (2010) and outline what we take to be a more fruitful argumentative analysis of the straw man fallacy, by combining PD and RT into a consistent framework. The crucial element of our proposal is an elaborate cognitive account of the pragmatic mechanisms involved in the hearer’s interpretation of fallacious straw man critiques. Finally, section 4 will illustrate the working of our proposal through the analysis of an example taken from Swiss political discourse. We will try to detail the subtleties of a naïve, speaker-to-speaker interpretation, consisting of context-bound and implicit meanings and conclude on the relationship between understanding and accepting. In plain words, we will be concerned with the relationship between what a speaker means and what the addressee ends up accepting.

2. Argumentative fallacies and cognitive heuristics

2.1 Heuristics and biases in Argumentation Theory

The idea of gaining a fuller account of fallacies by “bring[ing] the normative dimension better into relation with the psychological dimension” (Walton 2010: 160) is not new, and indeed it would be surprising if it were. Fallacies, in the end, while being discursively defined as misuses of argumentation, can also be regarded from a cognitive perspective as bias-generated errors in inferential processes of reasoning and judgment that we make ourselves (cf.

Correia 2011) or unwittingly accept when others make them. Such cognitive “errors” in our view have to be regarded as information-processing paths yielding outputs that would have been different – in particular for the epistemic sake of the cognitive system – in case the individual had been led to mobilise critical information. The nature of the error does therefore not lie in an inherently flawed inferential procedure but rather in the reliance on a heuristic procedure solicited in a situation that would be more advantageously dealt with by carefully considering additional critical information. Under this view, producing or failing to spot a fallacy is not a matter of doing something inherently “wrong”, but rather the result of failing to select the most optimal strategy given the circumstances. Whether this situation was deliberately aimed at by the speaker or not has no incidence on the cognitive characteristics of the fallacious move in terms of its processing. A genuinely cognitive account of fallacies should consequently not only explain how these biases work, but also specify the conditions under which they operate and turn out to be argumentatively and epistemically disadvantageous.³ We believe that part of the answer to these questions follows from the consequences of one of the defining features of fallacies, namely their covert nature. Fallacies are rhetorically effective, we contend, because they are not recognised as such, which is a property they share with deception (see Maillat & Oswald 2009, 2011, Oswald 2011, Oswald forth.).

Despite the manifest relevance of cognitive insights for the study of argumentation, research on cognitive aspects of reasoning (and by extension those of argumentation) has traditionally stayed within the bounds of cognitive psychology since Wason’s seminal works on logical reasoning in the 1960’s (Wason 1960, 1966) and Kahneman & Tversky’s pioneering work

3 A reviewer rightly pointed out that the notion of argumentative error should be carefully dealt with. It makes sense to speak of error only when a norm against which specific phenomena can be measured exists. The normative component of our account precisely allows us to judge whether arguments are fallacious or not, i.e., whether they amount to errors or not. On the cognitive side, the type of cognitive processing fallacies induce is not erroneous *per se*, to the extent that heuristics habitually perform well and yield good enough results, but suboptimal if we consider that the most cautious, reasonable and thus the most advantageous response in argumentation is the adoption of critical and reasonable doubt.

on cognitive heuristics (Kahneman & Tversky 1972, 1973).⁴ Yet, some attempts to incorporate the findings of this body of research have been made in the study of argumentation (e.g., Jackson 1996, Walton 2010), and it is to the discussion of these that we now turn.⁵

Jackson (1996) argues that the treatment of fallacies delineated by the question of “what is wrong with certain patterns of argument” should be complemented by “serious attention to what is persuasive about these patterns – why, given that they are defective, they often gain assent” (Jackson 1996: 101). She develops an account of the persuasiveness of fallacies from the perspective of her concept of conversational argument, by making use of empirical results in persuasion studies and social psychology. According to her, much of ordinary conversation is governed by “a broad and unremarked presumption of acceptability” (1996: 111) directed towards our interlocutors’ discourse. Jackson argues that we simply tend to accept what others claim, unless there are some important overriding reasons that lead us to challenge a given claim, such as inconsistency of the claim with what we already believe, or suspicions about the speaker’s trustworthiness and competence (see below for an elaboration on this very point). Motivated by such concerns, we take to examine the claim and its justification. The depth of the examination depends on factors such as our interest in the issue and our prior knowledge of it. Importantly, we stop our examination of (potentially fallacious) argumentation “whenever the cost of careful thinking exceeds what [we] think [...] the issue is worth” (Jackson 1996: 104) – in other words, we stop critical examination of a claim and its justification once we no longer expect it to be relevant to us. For such reasons, we tend to economise our assessment of argumentation by deploying simplified heuristics, or shortcuts of reasoning, that are often fast and frugal but not deeply critical (see Tversky & Kahneman 1974). As a result, we may be misled into accepting as sound a piece of argumentation that would end up being regarded as fallacious under closer scrutiny. In this sense, fallacies are treacherous because they are

4 See also Evans (2004), Evans & Frankish (2009), Gigerenzer *et al.* (2011) for state-of-the-art contributions on the topic.

5 It should nevertheless be noted that the links between argumentation and cognition are becoming increasingly salient in contemporary scholarly publications (e.g. Mercier & Sperber 2009, 2011, Mercier 2012, Bardone 2011, Maillat 2013, Correia 2011, Oswald 2011).

“incidental products of evaluation heuristics that can be given good defence as diagnostic tools” (Jackson 1996: 101). Jackson illustrates the working of her concepts by analysing fallacious appeals to authority and invalid forms of categorical syllogisms.

We find Jackson’s account very useful in terms of its reliance on a cognitive psychological body of research, but not satisfactorily elaborated for two main reasons: i) her particular conception of fallacies seems to miss a crucial normative dimension and ii) her account can be further grounded in light of recent findings in cognitive approaches to psychology and anthropology.

First, from our perspective, Jackson seems to have gone too far in claiming that fallacies lie, basically, in the eye of the beholder: “informal fallacies may describe not the materials presented by a speaker but the interpretive and reconstructive choices of the hearer” (Jackson 1996: 105). While we recognise that the second half of this characterisation is useful to address a research question concerned with a rhetorical account of *how* it is that fallacies may actually be effective, we believe that the first part of her formulation neglects the idea that fallacies are above all transgressive argumentative moves *realised in discourse* – i.e., they do lie in the “materials presented by a speaker” (*ibid.*). Accordingly, the study of fallacies equally needs to develop along a normative dimension, one which identifies clear criteria to discriminate sound from fallacious arguments. The identification of these criteria, we contend, constitutes a necessary step for the explanation of why fallacies in general can be defined as moves which violate (dialectical, logical) norms of reasonable argumentation, and why a given argument is fallacious in particular; this seems to be missing from Jackson’s account. Second, Jackson draws on the results of persuasion studies which are based on behavioural data but which do not explain exactly how the cognitive heuristics responsible for the success of fallacious arguments may actually work. We will turn to discuss research that frontally addresses this last issue in section 2.2.

Similarly to Jackson, Walton (2010) explores the possibility of elucidating the deceptive character of many informal fallacies of argumentation in terms

of their connection to cognitive heuristics.⁶ He does so by relying on the concepts developed in his theory of argumentation schemes. A heuristic, thus, is “a mediating concept between the notion of fallacy and defeasible argumentation scheme” (Walton 2010: 160). To explain this mediating function, Walton introduces the notion of a *parascheme*: “a device that can be used to represent the structure of a heuristic as a speedy form of inference that instinctively jumps to a conclusion and is commonly used to make decisions” (Walton 2010: 163). Paraschemes, which are schematic discursive devices used to represent the structure of argumentative heuristics, are argumentation schemes stripped of all the “prerequisites” that make a conclusion reached by a proper argumentation scheme a presumptively reasonable result of informal reasoning. These omitted prerequisites include additional premises, assumptions, rebutted exceptions, successfully answered critical questions, etc. Therefore, by relying on mere paraschemes, heuristics generate simplified and crude forms of argument prone to various errors and abuses that result in fallacies. Yet, these simplified forms are often astonishingly effective methods of reaching a conclusion, and thus they may “seem valid, but are not” (Walton 2010: 160, echoing the “standard” definition of fallacies – see Hamblin 1970: 12). According to Walton, here lies the treacherous character of many fallacies, especially those that are clearly based on the incorrect use of an argumentation scheme: *argumentum ad verecundiam* (argument from expert opinion), *argumentum ad ignorantiam* (argument from lack of knowledge), or *argumentum ad baculum* (argument from threat or bad consequences).

Walton’s account posits the existence of argumentative heuristics but does not draw on recent cognitive psychological research in support of his views. As a consequence, although the idea is very much appealing, it remains to be seen whether these dedicated paraschemes have cognitive plausibility at all, or if they rely on more general heuristics that may be used to attend to specific types of arguments when necessary. In other words, the specific role and mode of intervention of these heuristics is still unclear, from a cognitive perspective, in Walton’s account.

6 Walton seems to be unaware of Jackson’s research, or at least so is suggested by the absence of Jackson’s work in his discussion of the literature and in his reference list.

2.2 *Epistemic vigilance: cognitive insights into argument processing*

Jackson conceives of fallacies as “failed diagnostic strategies” (1996: 111). Although she presents this conception as an attempt to characterise what fallacies are, we believe, as pointed out above, that this provides only one half of a comprehensive account of fallacious argumentation, as the focus is here on how fallacies work from a cognitive perspective, and not on what fallacies are, *compared to sound arguments*. Still, we think that the direction of research she delineates – which is more in tune with rhetorical concerns such as argumentative effectiveness –⁷ is worth pursuing by mobilising insights from contemporary cognitive science. This sub-section will accordingly deal with cognitive constraints that may influence the persuasive success of fallacious argumentation.

There are indeed nowadays serious and promising grounds to start examining, within the framework of argumentation theory, Jackson’s characterisation of fallacies as failed diagnostics and assessing how they work from a cognitive perspective; for this we need to consider the cognitive constraints that can (mis)lead argument processing by recipients so as to drive their attention away from the fallacious nature of the argument. What we are therefore concerned with here is how fallacies manage to prevent their recipients from engaging in any sort of critical thinking with respect to the issue at stake. What in other academic circles is referred to as critical thinking has very recently started to attract attention in cognitive psychology and anthropology through the works of Sperber and his colleagues, who define our critical abilities in terms of epistemic vigilance (see Sperber *et al.* 2010 and its translation in this volume).

Sperber *et al.* posit that humans “have a suite of cognitive mechanisms for epistemic vigilance, targeted at the risk of being misinformed by others” (2010: 359). These cognitive filters are taken to monitor incoming information at all times, calibrating trust in its source while simultaneously

7 The notion of argumentative effectiveness will be here understood as an argument’s ability to bring about the desired perlocutionary effect in the addressee (i.e., being convinced, persuaded, etc.). For this reason, whether an argument is sound or fallacious is irrelevant to its qualification in terms of effectiveness. We also note that in this chapter the terms “effective” and “successful” will be used interchangeably to qualify arguments that have fulfilled this perlocutionary purpose.

assessing message consistency. The two dimensions (speaker reliability / trustworthiness and message consistency – both also mentioned by Jackson) are extremely relevant in particular when we consider the case of argumentative discourse, which is verbally manifested in communication as sets of propositions exhibiting a link of justification. Both sound and fallacious arguments should therefore be expected to try to satisfy epistemic vigilance requirements in order to promote their conclusions' perceived acceptability. Along these lines, it seems reasonable to assume that recipients of argumentation take into account at least the following two dimensions before they reach the stage where they are convinced (or unconvinced): i) the amount of trust they credit the speaker with and ii) the quality – in terms of consistency – of the information that is brought to their attention. The twofold idea behind epistemic vigilance, to put it crudely, is obviously that we tend to be more convinced by people we trust and messages we find consistent and truthful than by dodgy people and ill-evidenced or dubious, perhaps contradictory, messages. Fallacies take every chance to secure an apparent display of the characteristics of sound and acceptable arguments (i.e., trustworthiness and consistency). An inquiry into this phenomenon accordingly requires that we pay closer attention at trust and message consistency in communication.

2.2.1 Assessing trust (and exploiting the way we assess trust)

How can we defend, from a cognitive perspective, the assumption that fallacies may strive to remain undetected by exploiting trust? This first type of constraint could be taken to directly follow from what some philosophers define as a natural tendency not to be systematically sceptical in communicative settings, or, in other words, a tendency to regard as true – and believable – what others tell us (which is also what Jackson has in mind). This claim is fairly widespread in philosophical circles and can find its roots in Thomas Reid's conception of epistemic trust as a default disposition (cf. Reid 2000 [1794]). It is echoed nowadays in Ruth Millikan's work, who considers communication to be some kind of trustworthy perception by proxy (cf. Millikan 1987) and in Burge's take on communication, who considers in his *Acceptance principle* that we "accept as true something that is presented as true and that is intelligible [to us] unless there are stronger reasons not to do so" (Burge 1993: 467). Within contemporary pragmatics,

this very idea also underlies Grice's conversational maxim of quality, which enjoins conversational participants to provide information they consider themselves true and to refrain from contributing information they lack evidence for (see Grice 1989: 27).

This is obviously not to say that human beings are systematically and hopelessly gullible creatures; rather, this is a characterisation of a tendency to believe *in the absence of further evidence to the contrary*. The story nevertheless seems to be more complex, because there are grounds to assume that absence of further evidence to the contrary is the result of some sort of cognitive filtering mechanism.

Sperber *et al.* contend in this respect that what seems at first hand to correspond to default gullibility is not tantamount to blind trust; trust is not there by default as a precondition of communication, but it is mediated by epistemic vigilance filters; according to them, “[w]e could not be mutually trustful *unless* we were mutually vigilant” (Sperber *et al.* 2010: 364). Epistemic vigilance so described is thus partly directed at assessing whether we can trust the person who is informing or trying to convince us of something. These mechanisms are unconscious, to a large extent intuitive, and are supposed to alert us whenever there are grounds to suspect malevolent or deceptive intentions on behalf of the speaker. In other words, they are responsible, among other things, for checking whether there is available evidence not to trust a speaker.

Now, from the perspective of fallacious argumentation, it is reasonable to assume that a “successful” fallacy will manage to satisfy expectations of epistemic vigilance, so that no “red flag” is raised in the cognitive system as it is processed. And as far as monitoring the trustworthiness of the source is concerned, the fallacious argument will consequently precisely need to display every assurance that the speaker *is* trustworthy. In information-processing terms, considering someone to be trustworthy has important consequences on the epistemic features of whatever this person utters: crucially, we will deem what trustworthy people tell us to be epistemically strong – i.e., to reliably enhance our knowledge of the world. Managing to “discursively” establish a speaker's trustworthiness can therefore prove to be a way of managing to convey epistemically strong information, as long as it is said to stem from such a speaker. In argumentation parlance, we are therefore tackling speaker's *ethos*, namely how people perceive her in terms

of competence, benevolence, expertise, reliability, trustworthiness, etc. If we look at the traditional list of fallacies, the so-called “gang of eighteen” (see Woods 2004), one of them in particular seems to be exactly tailored to achieve this effect: the argument *ad verecundiam*, or appeal to the authority of expert judgment (cf. Walton 1989, Pilgram 2012).⁸ By launching an appeal to a (secretly) inadequate and irrelevant expert, the speaker is expecting the claim brought forth to inherit its epistemic strength from the figure she presents as an expert. If the appeal is successful, it means that the proposition associated with it enjoys a heightened epistemic status; what is more, if the addressee fails to see that the expert figure mentioned is inadequate, epistemic vigilance requirements might even be fulfilled, leaving the fallacy unthreatened as no reasons to question its validity have arisen.

2.2.2 *Assessing message consistency (and exploiting the way we assess message consistency)*

Alongside vigilance towards the source, epistemic filters are also vigilant towards the content of the communicated message. This means that we are naturally equipped to identify inconsistencies when the messages we process contain them. It is for instance straightforwardly unproblematic to perceive the inconsistency of contradictory or paradoxical sets of information (e.g., there is no particular difficulty in seeing the problem with the following formulation “Buenos Aires is the capital of Argentina and Buenos Aires is not the capital of Argentina”), to reject information that is blatantly false (e.g. “ $2 + 2 = 22$ ”), or to recognise that certain conclusions do not follow from certain premises (e.g. “All Dutchmen are Europeans. Laszlo is European. Therefore, Laszlo is Dutch”).

Verbally communicated information unfortunately does not always come in such “simple” forms. Complex argumentative operations, vague formulations, different sorts of propositions requiring sophisticated (meta)representations (about states of affairs, desirable states of affairs,

8 Recently, Pilgram (2012) distinguished (in a pragma-dialectical analysis) between an argument *by* authority and *from* authority. An argument *by* authority is an argument in which the arguer refers to her own authority to support a (sub)standpoint (such as when a doctor refers to her long clinical expertise in arguing for a given treatment). An argument *from* authority consists in an appeal to an expert external to the present discussion in support of a (sub)standpoint – a case we refer to.

mental states, etc.), requirements of field-specific background knowledge and further types of constraints might well complicate the task of being epistemically vigilant. Just because we are able to spot blatant inconsistencies quite effortlessly does not mean that we are unfailingly good at spotting any inconsistency. Among other parameters affecting consistency assessment, motivation, attention and relevance might act as powerful constraints.

Mercier & Sperber (2009, 2011) have devoted specific attention to one particular aspect of epistemic vigilance, namely argument-processing. Their argumentative theory of reasoning posits the existence of a dedicated argumentative module whose role is to take a claim and information relevant to its evaluation as input and to yield a judgment about the relationship between premises and conclusion, i.e. a judgment about the argument's validity or acceptability. When it is first solicited, the argumentative module delivers an *intuition* about the argument's validity. This is an output we are not typically aware of: at the end of the process, we accept or reject the claim without necessarily being aware of the reasons we have for doing so. However, this does not mean that these reasons are not there; they simply exist without us being aware of them. Nevertheless, when required, the argumentative module can disembed the conclusion of the argument to keep on processing it reflectively, thereby consciously attending to the reasons why the argument should be accepted in the first place. This is reasoning proper.⁹

This line of research is evidently of great interest to argumentation theory; however, what we would like to focus on here, while intimately related to the argumentative module, does not directly concern its internal machinery. We will instead be concerned with what the argumentative module takes as input. To the extent that assessing message consistency requires processing verbal information, it is reasonable to assume that "comprehension of the content communicated by an utterance is a precondition for its acceptance" (Sperber *et al.* 2010: 367). In any argumentative exchange, recipients indeed need, at least minimally, to carry out two cognitive tasks: they need to understand the content of both the premises and the conclusion (i.e., they need to understand the meaning of the speaker's complete argumentation) and they need to be able to produce some sort of evaluation – even if it remains

9 See Mercier & Sperber (2009, 2011) and Mercier (2012) for an in-depth discussion.

intuitive at first – about whether the conclusion follows from its premises. The idea we want to develop here, with the straw man fallacy as a case in point (see sections 3 and 4), is that mechanisms of comprehension can also be constrained, thereby affecting the subsequent validity judgment yielded by the argumentative module. In particular, we contend that many fallacies do not operate on the justificatory link *per se* but rather on the (non)selection of critical information. Following Maillat & Oswald's (2009, 2011) model of the Contextual Selection Constraint (henceforth CSC), we will assume that language users who are processing arguments in real time and on the spot may be misled into selecting information sets that are irrelevant to the evaluation of an argument – and in some cases even detrimental to the latter.

The CSC holds that in principle the meaning derivation procedure can be constrained along two dimensions corresponding to Sperber & Wilson's extent conditions of relevance (1995: 125), namely processing effort and cognitive effect. Processing effort translates into the amount of resources it takes to represent a given assumption and is consequently determined by information accessibility: the more accessible the assumption, the less effort it will take to process, and the more chances it has of being kept in the meaning derivation procedure. Cognitive effects are defined in epistemic terms by Sperber and Wilson: the assumptions that will provide the most reliable information about the world (in terms of addition of new information, or revision and suppression of old information) will be the most relevant to the cognitive system. Those assumptions that achieve the best ratio between effort and effect are deemed the most relevant. According to Relevance Theory, maximisation of relevance so conceived is a general feature of cognition; as a consequence, any selection of information (be it explicit information contained in a verbal stimulus or contextual information an addressee is responsible to mobilise on his own) will be governed by this mechanism.

Accordingly, discursive strategies that can be exploited to manage the effectiveness of fallacious arguments can either *strengthen* (that is, make more accessible in terms of processing effort or epistemically stronger in terms of cognitive effects) information sets which are devoid of critical information or *weaken* (that is, make less accessible or epistemically weaker) information sets which are loaded with critical information and which, if represented, would tentatively defeat the fallacious move by attracting attention on its

fallaciousness.¹⁰ The key point here is that these are constraints operating on *meaning*, to the extent that they manipulate both the accessibility and the strength of the information sets required to understand the message, which are the two extent conditions for reaching a relevant interpretation of a message. In view of that, we will show that the output of the comprehension procedure can be constrained and that this might have crucial implications for the overall judgment of argument validity or fallaciousness the recipient will be led to formulate (see the example in section 4 for a discussion).

2.2.3 Cognitive heuristics

A final note before we move on to the case of the straw man fallacy is in order. Our description of argument processing in light of recent work in cognitive psychology and on the study of verbal deception so far has highlighted how the way we process information can be constrained so as to influence the persuasiveness of verbal messages. How does that relate to cognitive heuristics?

Heuristics are traditionally defined as rules of thumb that help us take decisions, make judgments, and that more generally assist us in dealing with information in a cost-effective way. They are fast and frugal cognitive procedures that usually yield satisfactory results without requiring excessive amounts of cognitive resources:

Heuristics are frugal – that is, they ignore part of the information. Unlike statistical optimization procedures, heuristics do not try to optimize (i.e., find the best solution), but rather satisfice (i.e., find a good-enough solution). Calculating the maximum of a function is a form of optimizing; choosing the first option that exceeds an aspiration level is a form of satisficing. (Gigerenzer 2008: 20)

Traditionally, heuristics are associated with intuitive cognitive procedures, as opposed to effortful reflective ones. In this sense, both the comprehension

10 We should also mention here that the exploitation of these parameters is not an exclusive property of fallacies: sound arguments (and all discourse, actually), to the extent that they are comprehended following the same processing mechanisms that fallacious ones, are also subjected to these parameters. The difference between both is that fallacies block access to critical resources, while sound arguments in principle should foster it. The specificity of fallacies is thus to be found in their ability to obfuscate critical information.

module and the argumentative module may be considered to involve heuristics, to some extent at least.

We cannot help understanding an audible message that was uttered in our mother tongue (unless we block our ears or unless our attention is completely and consciously focused on something else); our interpretation of speaker meaning is most of the time unconscious and above all, we usually intuitively take what we understood from a verbal message to be reliable – in the sense that we do not engage in systematic doubt about what we have understood. Moreover, comprehension is fallible, to the extent that misunderstandings may occur. These usually result from a mismatch between linguistically decoded and/or contextual information selected by the addressee and linguistically encoded and/or contextual information selected by the speaker as she was planning and formulating her utterance. As fallibility is among the core features of heuristics, comprehension can be described as a heuristic. The argumentative module is also of a heuristic nature because it yields an intuitive inference, which is also fallible. Sometimes we are convinced by ill-evidenced claims, and further (reflective) examination allows us to identify the flaw. The effectiveness of fallacious arguments can furthermore be taken as evidence for the heuristic nature of default argument processing.

In this section we have tried to characterise more precisely what it means for a fallacy to tap on cognitive heuristics. We have focused in particular on how fallacies can go past epistemic vigilance filters, notably by trying to display expertise and consistency. We interpreted this in terms of strengthening and weakening strategies, and we now turn to discuss these mechanisms through our account of the straw man fallacy, which we then illustrate by an example in section 4.

3. Accounting for the straw man fallacy

3.1 *Walton's account*

In his paper on heuristics, Walton lists the straw man fallacy among twelve informal fallacies that “need to be analysed with defeasible argumentation

schemes” (2010: 174-175). The straw man is thus an abuse of the “argument from commitment” as defined by Walton, Reed and Macagno (2008: 335). In a contemporary paper, however, Walton takes a rather different, if not the opposite, stance:

With many of the informal fallacies, the problem of fallacy identification, analysis, and evaluation is made easier by the fact that the fallacy is closely related to a known argumentation scheme [...]. This resource is not available, however, in the cases of wrenching from context and straw man. These fallacies are more purely dialectical in nature. They basically arise from or result in the misrepresentation of one party’s position by another party with whom the first party is engaged in a dialogue. (Walton & Macagno 2010: 303-304)

We agree with the latter formulation. In our view (Lewiński 2011, Lewiński & Oswald 2013), the straw man cannot be a violation of an argument scheme, because strictly speaking the straw man is not an argument in the first place. Rather, it is an attack on someone else’s argument; this is why, for instance, a real question (not a rhetorical one) can convey a straw man but not an argument. In terms of the pragma-dialectical critical discussion, a straw man is a (fallacious) speech act performed by the critical antagonist, rather than by the argument-making protagonist. In our analysis this accounts for the “dialectical nature” of the straw man fallacy, as Walton and Macagno (2010) have put it. One way of dealing with Walton’s inconsistency in approaching the straw man is to treat the second formulation (Walton & Macagno 2010) as a misrepresentation of dialectical commitments in terms of the form of the fallacy, and the first formulation (Walton 2010) as a type of attack from commitment in terms of the function of the straw man. Only a clear grasp of both the form and the function of the straw man would define the fallacy, as argued by Lewiński (2011: 480). Walton, thus, would propose two mutually exclusive accounts of the straw man, since he actually speaks of two different aspects of the straw man that eventually can be seen as complementary, rather than exclusive. Such an interpretation, while highly charitable, is not explicitly put forth by Walton himself, and thus may not accurately capture his take on the fallacy. Noticeably, however, Walton (and Macagno) offer no explanation, in terms of argumentation schemes or *paraschemes*, of how straw man attacks are deceptive and may pass for good, reasonable forms of argumentative criticism.

In the following, we will propose a different take on the straw man fallacy that clearly relates fallacy evaluation to precisely described cognitive mechanisms. Our contribution is thus novel not in its general aim (since both Walton and Jackson have undertaken a similar line of inquiry), but rather in the exact way we envisage a fruitful cross-fertilisation of recent developments in normative theories of argumentation and cognitive studies. From the perspective of our goals, even if Walton and Jackson both link fallacies and heuristics, they are not attentive enough to cognitive mechanisms underlying the deceptiveness of many fallacies. Moreover, they do not offer an applicable normative account of fallacies that supports a consistent analysis of the straw man: while Jackson chooses a decidedly descriptive path, Walton seems to offer two incompatible accounts of how to best analyse a straw man as a fallacy. The advantage of our work is thus, first, a higher level of precision in both its normative and cognitive aspects, and, second, a focus on the detailed working of one particular fallacy – the straw man – that seems somewhat elusive in other accounts.

3.2 An integrated normative and cognitive pragmatic account of the straw man fallacy

In terms of the integrated pragma-dialectical theory, a straw man, similarly to any other fallacious argumentative move, is conceptualised as a “derailment of strategic manoeuvring” (van Eemeren & Houtlosser 2003). *Strategic manoeuvring* is defined as a continuous balancing between two competing yet reconcilable goals in argumentation. On the one hand, in order to resolve a difference of opinion in the process of genuine critical testing of their opinions, arguers should meet dialectical requirements of reasonableness embodied in the rules of the PD model of a critical discussion (van Eemeren & Grootendorst 2004: Chs. 6 and 8). On the other hand, in attempts to have their position accepted by the antagonist, arguers may take advantage of numerous rhetorical techniques. While manoeuvring aimed at reconciling the reasonable with the opportune in argumentation is perfectly possible, the strife for persuasive success may also lead to the abandonment of standards of reasonableness. Whenever the latter happens, pragma-dialecticians speak of a derailment of strategic manoeuvring that by definition amounts to

committing a fallacy, because a rule for a reasonable critical discussion is violated.

The basic PD understanding of the straw man fallacy is thus that of a violation of rule 3 for a critical discussion: *A party's attack on a standpoint must relate to the standpoint that has indeed been advanced by the other party* (van Eemeren & Grootendorst 1992: 124-131).¹¹ Contravening this rule and hence attacking a straw man seriously hinders critical testing. The antagonist who surreptitiously misrepresents the protagonist's opinions seems to be involved in genuine critical testing of these opinions, but in fact attempts no more than a made-up falsification aimed at rhetorical victory (traditionally termed a "sophistical refutation"). As a result, the possibility of reaching a reasonable resolution of the entire dispute on the merits is seriously undermined.

Yet, even though the possible rhetorical advantage that can be gained by committing fallacies is seen in PD as a driving force behind much of fallacious argumentative discourse, this theory does not aim at expounding how such fallacious persuasive appeals actually work.

We have consequently tried to establish elsewhere (Lewiński & Oswald 2013) the grounds for the complementarity of cognitive pragmatics (and Relevance Theory specifically) and PD in an exhaustive treatment of the straw man fallacy. We hypothesised that accounting for the straw man fallacy is a matter of interpretation of argumentative discourse *on two distinct levels*. The first level of interpretation, which is the normative level of argumentation analysis, is concerned with providing identifying criteria for the straw man fallacy. The second level of interpretation, dedicated to explain why straw men may be effective, considers mechanisms of comprehension and how these constrain what recipients of the straw man fallacy understand from the message. Specifically, we claim that a cognitive pragmatic model of utterance interpretation is ideally equipped to explain how

11 We note here that the rules of the critical discussion, as identified in PD, are not comparable to conversational maxims (cf. Grice 1989), whose flouting is instrumental to the derivation of meaning (in terms of implicature identification): pragma-dialectical rules ensure critically reasonable argumentative discussions and the recognition that these have been violated in a given discourse allows the analyst to pass a judgement on the fallaciousness of the argument. In other words, the rules of the critical discussion do not have anything to do with meaning.

misrepresentations involved in a straw man attack may pass for genuine and faithful representations.

We thus first stipulated certain general contextual conditions under which an analyst can judge a given interpretation of the protagonist's position a straw man. We formulated the first basic criterion of identification for the straw man fallacy in terms of pragmatic plausibility: as long as the antagonist follows certain contextual possibilities in deriving speaker meaning and thus as long as she stays within the bounds of a disagreement space of a given utterance, she cannot be seen as committing a straw man.¹² The second crucial criterion for normative interpretation is interpretative charity. Charity of interpretation – i.e., the choice of an interpretation that is most beneficial to the arguer – is typically advised to arguers and analysts alike in cases of interpretative doubt (see Lewiński 2011, 2012). However, rather than perceiving it as a rule of reasonable argumentation, we take charity between arguers to be a rhetorical choice which can be made one way or another. The antagonist who is capable of producing compelling criticisms against the protagonist's position may opt for a very charitable interpretation of the protagonist's argumentation. Yet, she can also opt for an uncharitable interpretation – and in principle there is nothing wrong with such an interpretation as long as it is pragmatically plausible. All the same, we contend that similarly to pragmatic plausibility the criterion of interpretative charity between arguers is inherently contextual. In some contexts that can be termed *constructive*, arguers are expected, indeed required, to be more charitable than in *critical* contexts. Similarly, some contexts call for heightened *precision* (thus limiting the scope of plausible interpretations), while others allow for *loose* interpretations where plausibility is much extended (see Lewiński 2011: 490ff.).

Even if satisfactorily laid out, however, such conditions pertain to an analyst's "normative interpretation" and evaluation based on a specific theory of argumentation. Therefore, such inquiry does not answer the question of how ordinary addressees, in the process of pre-theoretical, natural comprehension (or "naïve interpretation") are misled into

12 The notion of disagreement space denotes the "entire complex of reconstructible commitments" and consequently defines "a structured set of opportunities for argument" (van Eemeren *et al* 1993: 95). Pragmatically speaking, the disagreement space consists of all contextually plausible interpretations of a given utterance.

interpreting argumentative discourse in a way that will leave the straw man unidentified. Building on a cognitive pragmatic model of information-selection, we argued along the lines exposed above in section 2.2.2 that straw men fallacies effectively manage to constrain meaning so that misattributions of commitment pass for correct attributions of commitment. This process can unfold along the abovementioned two types of constraining strategies: the relevance of potential clues revealing the misattribution of commitment can be weakened, while evidence of a sound attribution of commitment is strengthened.

Building on the complementarity of the two pragmatic approaches used, we gave a fuller account of the working of the straw man fallacy, in terms of both its dialectical incorrectness and rhetorical treacherousness; in what follows we turn to discussing a real example to illustrate our account.

4. Exhuming the straw man: a concrete example

During one of the meetings of the Swiss National Council in December 2008, Carlo Sommaruga, a representative for the canton of Geneva, officially submits and motivates a motion meant to legally rehabilitate the Swiss fighters of the International Brigades, who fought in Spain against Franco's dictatorship in the 1930s.¹³ Most of these fighters, upon their return to Switzerland, were prosecuted on the grounds that the Swiss military criminal code does not authorise its citizens to serve in foreign armies, and most of them were sentenced to prison (with sentences ranging from 15 days to 4 years of imprisonment, together with a denial of their civic rights). Oskar Freysinger, another member of the lower house of the Swiss parliament, opposes this motion with the following argumentation:

13 The official text of the motion (in French) can be found online in the archives of the Swiss parliament: http://www.parlament.ch/ab/frameset/f/n/4806/284915/f_n_4806_284915_285114.htm (Last accessed 01.05.2013).

Mr. Sommaruga, you are painting a very romantic portrait of the International Brigades. However, did you know that George Orwell, the very well-known author who fought in their ranks, very quickly noted that these people aimed at the establishment of a Stalinist regime and that their violence was as serious as the one perpetrated by the fascists? So, do you want to rehabilitate people who intended to fight a totalitarian regime in order to replace it with another totalitarian regime?¹⁴

The last sentence of Freysinger's intervention, which is realised as a rhetorical question, can be taken to convey both the standpoint of his argumentation and a simple justification for it; we give three reasons in support of this reading. First, in contemporary democratic countries such as Switzerland, no reasonable person would indeed support a totalitarian regime, based on the widespread idea that totalitarianisms are undesirable political systems. Second, the information provided before the question and which relates to the International Brigades' violence is presented by Freysinger as evidence in support of their non-rehabilitation. Finally, Freysinger's contribution is immediately consecutive to Sommaruga's exposition of the motion: in the broader context of this particular activity type, Freysinger's question functions as an opposition to what was proposed.¹⁵ For these reasons, we take the main standpoint of the argument to be the following:¹⁶

1. Standpoint: The Swiss fighters of the International Brigades should not be legally rehabilitated.

Several arguments are presented in support of the claim. Looking at the rhetorical question, it appears that according to the speaker these fighters were seeking to replace a totalitarian regime by another totalitarian regime,

14 Here is the original text in French: "*Monsieur Sommaruga, vous nous dressez un portrait très romantique des Brigades internationales. Or, savez-vous que George Orwell, auteur très connu, qui a combattu dans leurs rangs, a très vite remarqué que ces gens visaient à l'établissement d'un régime stalinien et que leurs violences étaient tout aussi graves que celles que commettaient les fascistes? Alors voulez-vous réhabiliter des gens qui cherchaient à combattre un régime totalitaire afin de le remplacer par un autre totalitarisme?*".

15 Note that exactly because Freysinger's standpoint is the contradictory of Sommaruga's proposal, the difference of opinion is "mixed" (see van Eemeren & Grootendorst 2004: 119-120): Freysinger builds his argumentation to both support his position and criticise Sommaruga. This situation warrants an analysis of his arguments in terms of the straw man, as defined above.

16 We are following the PD notation of the structure of arguments in an analytic overview. See van Eemeren *et al.* (1993).

and this, combined with an implicit premise pointing to the undesirability of this state of affairs, gives us a first argument in support of (1):

- 1.1a Explicit premise: The International Brigades wanted to replace a totalitarian regime by another totalitarian regime.
- (1.1a') Implicit premise: We should not legally rehabilitate people who want to replace a totalitarian regime by another totalitarian regime.

What precedes the rhetorical question functions as evidence for its content: Freysinger is using what he knows about the International Brigades to legitimise the crucial idea that the fighters of these brigades were promoters of a totalitarian regime. This means that the explicit premise (1.1a) becomes in our reconstruction of the argumentative chain the standpoint of a sub-argument. In what precedes, we do find at least one relevant required premise, which, combined with an additional implicit premise accessed through our encyclopaedic knowledge, allows us to reach (1.1a); to the extent that the conclusion from (1.1a.1) and (1.1a.1') is not exactly worded as (1.1a), we represent it in (1.1a*) as an intermediate step allowing us to spell out the link between (1.1a.1) and (1.1a.1') on the one hand and (1.1a) on the other:

- 1.1a.1 Explicit premise: These people aimed at the establishment of a Stalinist regime.
- (1.1a.1') Implicit premise: The Stalinist regime was totalitarian.
- 1.1a* Implicit standpoint: These people aimed at the establishment of a totalitarian regime.

This standpoint is not exactly formulated as the premise we identified in (1.1a); however, both (1.1a*) and (1.1a) have in common the idea that the International Brigades, according to Freysinger, wanted to bring a totalitarian regime to power. Moreover, they are both supported by the same premises that taken together satisfy the “logical minimum” (van Eemeren & Grootendorst 2004: 117-118) of argumentation. Next to this sub-argument, Freysinger also notes that the violent nature of the International Brigades is as serious as the violent actions committed by the fascists. This information becomes relevant by acquiring an argumentative function, though this time not on a subordinate level, but directly with respect to the main standpoint (1), as it contributes an additional reason not to legally rehabilitate the Swiss fighters. Here is the reconstruction of this particular argument with an

explicit formulation of the implicit premise required to support the standpoint:

- 1.1b Explicit premise: The violence of the International Brigades was as serious as the one perpetrated by the fascists.
- (1.1b*) Implicit premise: We should not legally rehabilitate those whose violent actions were as violent as the actions of the fascists.
- 1 Standpoint: The Swiss fighters of the International Brigades should not be legally rehabilitated.

Breaking the argumentation further down into its different parts, we notice that both (1.1a.1) and (1.1b) gain their epistemic strength from their contents being attributed to George Orwell, “the very well-known author”. Here, Freysinger is trying to make the two premises epistemically stronger by taking advantage of Orwell’s credit in its quality of direct witness of what was happening in the ranks of the International Brigades – after all, he was one of them and he was there, as asserted by Freysinger. It is not our purpose here to assess in detail whether this appeal to authority is fallacious or not, even though we think there is considerable room to debate the issue.¹⁷ It remains that the use of Orwell’s authority by the speaker intervenes as a way of establishing the acceptability of premise (1.1a.1) and premise (1.1b), which are part of two coordinated argumentations in support of the main standpoint (1).

A careful reconstruction and synthesis of the argument leads us to the following global paraphrase:

- 2. We should not legally rehabilitate the Swiss fighters of the International Brigades because they wanted to replace a totalitarian regime by another totalitarian regime (since, as noted by Orwell, they promoted a Stalinist

17 We have no conclusive evidence that the appeal to authority here is fallacious, but we note that Orwell’s credibility as an author does not give him the authority to judge if the International Brigades were violent or if they wanted to establish a totalitarian regime; furthermore, his taking part to the action of the International Brigades is not sufficient either to assess his credibility – we would need to know when he was part of them, his motivation for joining them, and his reasons for leaving them. Still, an *ad verecundiam* argument seems to be lurking here and it is probably intended by Freysinger to make (1.1a.1) and (1.1b) epistemically stronger than other assumptions. See also Herman (this volume) for a detailed discussion of arguments from authority.

regime) and because – and this is also noted by Orwell – their violence was as serious as the violence perpetrated by the fascists.

There is still one step missing in our analysis, and it has to do with the link that Freysinger creates between the International Brigades wanting to promote totalitarianism and the fact that they were promoting Stalinism. In other words, we are concerned here with the shift from premises (1.1a.1) and (1.1a.1') to the conclusion (1.1a*). While (1.1a.1) is, as we have seen, supported by the appeal to Orwell's authority, nothing is adduced in support of the implicit premise (1.1a.1'), which is nevertheless crucially needed to reach the conclusion (1.1a*), itself subsequently needed to reach the main standpoint of the argument (1). And this is precisely where we contend that the straw man is hidden. Unexpressed premise (1.1a.1') tells us that Stalinism was a totalitarian regime. Indeed, nowadays we know that such was the case, and that the concentration of power into one man's hands led the former Soviet Union to suffer the worst dictatorship of its history. Yet, is it fair to consider that the Swiss fighters of the International Brigades were supporting a totalitarian regime? Freysinger tells us that they supported Stalinism, which is true. But in the 1930s context, did supporting Stalinism straightforwardly amount to supporting a dictatorship? This is far from certain. It is only after the 1930s that the world gradually started to find out about the horrors of Stalinism; before that, few had witnessed them, and it would probably be more accurate to say that the International Brigades had as an ideal the promotion of a regime which, on paper, was the complete opposite of a dictatorship. Furthermore, in the 1930s, communism under Stalin had not yet become a totalitarian regime – so there was no way people could know it. Therefore, accusing the Swiss fighters of promoting a totalitarian regime is uncharitable – and furthermore anachronistic since they could not know that Stalinism was going to become such a regime. Had they known what Stalinism became, they would most likely never have supported it, since that is precisely the type of regime they were fighting at that time (i.e., Franco's fascist regime).

In this example, Freysinger is using a problematic piece of encyclopaedic knowledge in (1.1a.1') as an implicit premise to support the implicit conclusion (1.1a*), which is itself, as a misrepresentation of the Swiss fighters' motivations, used in the global argument in order to attack Sommaruga's main standpoint. The straw man in this example seems to be

multi-layered and complexly structured: its first step consists in misrepresenting the intentions of a group of people targeted by the motion, namely the fighters of the International Brigades; we have seen that the misrepresentation can be qualified as such because it is both uncharitable, but also pragmatically implausible by virtue of its anachronistic nature. It should be highlighted in this respect that the treacherousness of the argumentative move here does not lie in the explicit premise (1.1a.1), because the International Brigades indeed wanted a Stalinist regime. The problem is that in the 1930s context (1.1a.1') was not known, and therefore it should not legitimately be counted among something these fighters could reasonably be taken to believe.

But the actual straw man in the parliamentary discussion in question, instantiated here as an attempt to counter Sommaruga's standpoint, only becomes fully fledged as Freysinger formulates his rhetorical question. Through this question, Freysinger implies that Sommaruga endorses the (misattributed) views of the International Brigades about promoting Stalinism. The move consequently also results in a misattribution of commitment to Sommaruga: as Freysinger puts it, his opponent is guilty of intending to rehabilitate people who were no better than fascists. As noted above, such a criticism is not only uncharitable – which is expected of two deputies from opposing parties and thus contextually acceptable – but also implausible due to an anachronistic ascription of intentions: one could hardly be a supporter of totalitarian Stalinist methods before these methods became known or even existed. Therefore, Freysinger's critique of Sommaruga's motion does qualify as a straw man fallacy, even though of a rather subtle, pragmatic type. To sum up on the subtlety – and complexity – of the straw man in Freysinger's intervention: his misattribution of beliefs to the fighters of the International Brigades is not a straw man fallacy *per se*, to the extent that there is no argumentative discussion between him and these fighters. However, by misrepresenting their beliefs and by implicitly communicating that Sommaruga is liable to endorsing them through his motion, Freysinger makes as if these misrepresentations are also part of Sommaruga's actual commitments.

A closer look at the linguistic choices used by Freysinger in this complex argumentation will reveal how the relevance of critical information is weakened, while the fallacy's apparent relevance as a sound argument is at

the same time strengthened. First of all, the rhetorical question is directly addressed at Sommaruga in a vocative mode (“Do *you* want to rehabilitate...”). It would only be relevant to ask such a question if there were plausible reasons to consider that Sommaruga indeed had voiced the intention of rehabilitating people who promoted a totalitarian regime. Moreover, a rhetorical question, as pointed out by Snoeck Henkemans, “serves as a means to urge the addressee to act on his commitment and recognise that the standpoint that is being defended by the argument the addressee supposedly accepts, should now also be accepted” (Snoeck Henkemans 2009: 17); here Freysinger is (deceptively) trying to get Sommaruga to endorse the implications of one of his alleged previous commitments for argumentative purposes. The question under discussion is hence not at all whether the International Brigades promoted a totalitarian regime, but rather whether Sommaruga is prepared to accept that such was the case. By constraining the question’s focus onto Sommaruga’s commitments, Freysinger may be charged of attempting to direct his audience’s attention (including his opponent’s) on the attack rather than on the grounds of the attack. In other words, the rhetorical question *strengthens* the relevance of Sommaruga’s supposed mistake, because, contextually, that’s the very point of Freysinger’s intervention. As for the grounds on which the attack is launched, their relevance in the exchange is made less salient and is *ipso facto* weakened.

From the perspective of strengthening, an additional parameter appears to be of importance: the *ad verecundiam* we identified earlier is meant to strengthen the epistemic status of (1.1a.1) as it attempts to take advantage of Orwell’s presumed credibility. By resorting to this strategy, Freysinger intends to increase the chances of (1.1a.1) being accepted as a reliable piece of information. Then, once (1.1a.1) is accepted and after (1.1a*) is processed as an explicit content of the rhetorical question, the only way to understand the argumentative link, or, in other words, the relevance of (1.1a.1) with respect to (1.1a*) is to represent the implicit premise (1.1a.1’). In other words, in order to *understand* the link between (1.1a.1) and (1.1a*), we need to mobilise (1.1a.1’) – albeit only intuitively. And since (1.1a.1’) is necessary for comprehension purposes, it might be argued that it will be part of the meaning derivation procedure, thereby increasing its chances of being selected. Crucially, the use the audience is asked to (intuitively) make of

(1.1a.1') is not related to any kind of argumentative evaluation; it is instead presented as a condition for the meaning of Freysinger's contribution to be properly understood. We claim that this weakens the chances of (1.1a.1') getting "argumentatively" picked up. In other words, the addressee's cognitive environment is encouraged to consider that (1.1a.1') is relevant prominently with respect to the comprehension procedure, and not so much with respect to the argumentative procedure.¹⁸

More can be said on the particular status of (1.1a.1'). The second parameter affecting the salience of this first commitment misattribution is its locus of occurrence in the overall argumentation. (1.1a.1') is part of a subordinate argument which is used in support of one of the coordinated arguments (1.1a) adduced in support of the main standpoint in (1). The position of this misattribution is probably not fortuitous; since it is embedded in a sub-argument – and moreover not presented as a piece of information that needs to be called into question due to its prominent interpretative role –, it is not presented as decisive for the global argument. Again, this can be interpreted in terms of salience: for the whole argument to make sense, (1.1a.1') need not be consciously represented, which reduces the chances of the content being submitted to critical doubt. Put differently, (1.1a.1') is just a mere background to the main argumentation. As such, it is presented as if its weight was not pivotal in the overall strength of the case.

Summing up, we contend that the example of Freysinger's criticism is an interesting case of the straw man fallacy. It remains rather subtle and thus requires close critical attention within a framework of a context-sensitive,

18 The nature of (1.1a.1') in terms of meaning components can be further discussed. While it seems pretty clear, as argued by an anonymous reviewer, that (1.1a.1') qualifies as an implicated premise, there could be room to suspect that it is a *discursive* presupposition (distinct from classical semantic presuppositions, see Saussure, 2012, forth. and this volume for a discussion of the notion): its function, in addition to its argumentative role, would be to provide the necessary bits of information for the link between (1.1a.1) and (1.1a*) to plainly make sense. Interestingly, the contribution of presuppositions to meaning is subtle and occurs somewhat "backstage". Some have claimed that the contents of presupposition can accordingly be forced into communication while escaping the addressee's critical awareness (see Saussure 2012, and Polyzou 2013 for instance). We also take this to be an indication that the tentative discursive presuppositional status of (1.1a.1') might be processed by the comprehension module and fail to qualify as input for the argumentative module.

normative pragmatic theory, such as PD, to be clearly assessed as fallacious. But because of this subtlety, the example also vividly illustrates the different weakening and strengthening strategies employed in Freysinger's discourse that may affect the way the argument is processed by the audience of his attack. A first misattribution of commitment, whose argumentative potential is obscured by its importance in the meaning derivation procedure, is used to establish the grounds on which the primary straw man fallacy operates. In other words, comprehension has constrained argumentation. By implicitly leaving room for the audience to infer that Sommaruga's motion is an indication that he might endorse totalitarian views himself, Freysinger tries to push his opponent into defending himself, which might turn into a "cognitive priority", at the expense of a critical examination of the argument.

5. Conclusion

Accounting for the straw man fallacy is a clear example of the need for pragmatically-driven theoretical research. In this paper we have tried to show how a combination of two contemporary theories with strong pragmatic import may provide a grounded alternative to existing accounts by addressing both normative and cognitive questions.

Reliance on PD makes for the requirements of the question of fallacy identification; reliance on cognitive pragmatics but also on cognitive psychology and anthropology makes for the requirements of the question of fallacy effectiveness. It was one of the purposes of this paper to highlight the role that cognitive studies should be able to endorse within the framework of argumentation theory.

Underlying our exposition of the persuasive effectiveness of the straw man fallacy are questions of great cognitive significance; notably, we touched here upon the complex issue of the way verbal stimuli are processed, depending on their relevance for different cognitive tasks. The example discussed in section 4 has shown that the same representation (in this case the implicit premise (1.1a.1')) might be processed either by the comprehension module or by the argumentative module. Crucially, we

argued that in the example the interpretative path needed to be privileged in order for the fallacy not to be critically tested, that is, in order to block its processing by the argumentative module. It might be an indication that variations in communicative contexts might trigger competition between cognitive modules. This could well represent a new direction of research in the study of rhetorical effectiveness.

References

- Aikin, S. J. & Casey, J. (2011): "Straw men, weak men, and hollow men", *Argumentation* 25 (1), 87-105.
- Bardone, E. (2011): *Seeking chances. From biased rationality to distributed cognition*. Springer, Berlin/Heidelberg.
- Bizer, G., Kozak, S. & Holterman, L. (2009): "The persuasiveness of the straw man rhetorical technique", *Social Influence* 4 (3), 216-230.
- Burge, T. (1993): "Content preservation", *Philosophical Review* 101, 457-88.
- Correia, V. (2011): "Biases and fallacies: The role of motivated irrationality in fallacious reasoning", *Cogency* 3 (1), 107-126.
- Eemeren, F. van & Grootendorst, R. (2004): *A Systematic Theory of Argumentation: The Pragmadiadical Approach*. Cambridge University Press, Cambridge.
- (1992): *Argumentation, Communication, and Fallacies: A Pragmadiadical Perspective*. Lawrence Erlbaum, Hillsdale, NJ.
- Eemeren, F. van, Grootendorst, R., Jackson, S. & Jacobs, S. (1993): *Reconstructing argumentative discourse*. The University of Alabama Press, Tuscaloosa/London.
- Eemeren, F. van & Houtlosser, P. (2003): "Fallacies as derailments of strategic maneuvering: The *argumentum ad verecundiam*, a case in point", in Eemeren, F.H. van, Blair, J.-A., Willard, C.A., Snoeck Henkemans, F. (eds), *Proceedings of the Fifth Conference of the International Society for the Study of Argumentation*. SicSat, Amsterdam, 289-292.
- Evans, J.S.B.T. (2004): "Biases in deductive reasoning", in Pohl, R. (ed.) *Cognitive illusions. A handbook on fallacies and biases in thinking, judgement and memory*. Psychology Press, Hove/New York, 127-144.
- Evans, J.S.B.T. & Frankish, K. (eds) (2009): *In Two Minds*. Oxford University Press, New York.
- Gigerenzer, G. (2008): "Why Heuristics Work", *Perspectives on psychological science* 3 (1), 20-29.
- Gigerenzer, G., Hertwig, R. & Pachur, T. (eds) (2011): *Heuristics: The foundation of adaptive behavior*. Oxford University Press, New York.
- Grice, H. P. (1989): *Studies in the way of words*. Harvard University Press, Cambridge, MA.
- Hamblin, C. (1970). *Fallacies*. Methuen, London.

- Herman, T. (2014). "L'argument d'autorité: structure, définition et effets rhétoriques", in Herman, T. & Oswald, S. (éds), *Rhétorique et cognition*. Peter Lang, Berne.
- Jackson, S. (1996): "Fallacies and heuristics", in Bentham, J. van, Eemeren, F.H. van, Grootendorst, R. & Veltman, F. (eds), *Logic and argumentation*. Royal Netherlands Academy of Arts and Sciences, Amsterdam, 101-113.
- Kahneman, D. & Tversky, A. (1972): "Subjective probability: a judgement of representativeness", *Cognitive Psychology* 3, 430-454.
- Kahneman, D. & Tversky, A. (1973): "On the psychology of prediction", *Psychological Review* 80, 237-251.
- Laar, J. van (2008): "Room for maneuver when raising critical doubt", *Philosophy and Rhetoric* 41 (3), 195-211.
- Lewiński, M. (2011): "Towards a critique-friendly approach to the straw man fallacy evaluation", *Argumentation* 25 (2), 469-497.
- (2012): "The paradox of charity", *Informal Logic* 32 (4), 403-439.
- Lewiński, M. & Oswald, S. (2013): "When and how do we deal with straw men? A normative and cognitive pragmatic account", *Journal of Pragmatics* 59 (Part B), 164-177.
- Maillat, D. (2013): "Constraining context selection: on the pragmatic inevitability of manipulation", *Journal of Pragmatics* 59 (Part B), 190-199.
- Maillat, D. & Oswald, S. (2009): "Defining manipulative discourse: the pragmatics of cognitive illusions", *International Review of Pragmatics* 1 (2), 348-370.
- (2011): "Constraining context: a pragmatic account of cognitive manipulation", in Hart, C. (ed.), *Critical discourse studies in context and cognition*. John Benjamins, Amsterdam, 65-80.
- Mercier, H. (2012): "Looking for arguments", *Argumentation* 26 (3), 305-324.
- Mercier, H. & Sperber, D. (2009): "Intuitive and reflective inferences", in Evans, J.S.B.T. & Frankish, K. (eds), *In Two Minds*. Oxford University Press, New York, 149-170.
- (2011): "Why do humans reason? Arguments for an argumentative theory", *Behavioral and Brain Sciences* 34 (2), 57-74.
- Millikan, R. G. (1987): *Language, Thought and Other Categories*. MIT Press, Cambridge, MA.
- Oswald, S. (2011): "From interpretation to consent: Arguments, beliefs and meaning", *Discourse Studies* 13 (6), 806-814.
- Oswald, S. (forth.): "It is easy to miss something you are not looking for: A pragmatic account of covert communicative influence for (critical) discourse analysis", in Hart, C. & Cap, P. (eds), *Contemporary Studies in Critical Discourse Analysis*. Continuum, London.
- Pilgram, R. (2012): "Reasonableness of a doctor's argument by authority: A pragma-dialectical analysis of the specific soundness conditions", *Journal of Argumentation in Context* 1 (1), 33-50.
- Reid, Th. (2000) [1764]: *An Inquiry into the Human Mind and the Principles of Common Sense*. Penn State Press, University Park, PA.
- Saussure, L. de (2013): "Background relevance", *Journal of Pragmatics* 59 (Part B), 178-189.
- Saussure, L. de (2014). "Présuppositions discursives, assertion d'arrière-plan et persuasion", in Herman, T. & Oswald, S. (éds), *Rhétorique et cognition*. Peter Lang, Berne.
- Snoeck Henkemans, F. (2009): "Manoeuvring strategically with rhetorical questions", in van Eemeren, F.H. & Garssen, B. (eds), *Pondering on problems of argumentation*. Springer, New York, 15-23.

- Sperber, D., Clément, F., Heintz, C., Mascaro, O., Mercier, H., Origg, G. & Wilson, D. (2010): "Epistemic vigilance", *Mind and Language* 25 (4), 359-393.
- Sperber, D. & Wilson, D. (1995) [1986]: *Relevance. Communication and cognition*. Blackwell, Oxford.
- Talisse, R. & Aikin, S. F. (2006): "Two forms of the straw man", *Argumentation* 20 (3), 345-352.
- Tversky, A. & Kahneman, D. (1974): "Judgment under uncertainty: Heuristics and biases", *Science*, New Series 185 (4157), 1124-1131.
- Walton, D. (1996): "The straw man fallacy", in Benthem, J. van, Eemeren, F.H. van, Grootendorst, R., Veltman, F. (eds), *Logic and Argumentation*. Royal Netherlands Academy of Arts and Sciences, Amsterdam, 115-128.
- Walton, D. (2010): "Why fallacies appear to be better arguments than they are", *Informal Logic* 30 (2), 159-184.
- Walton, D., Reed, C. & Macagno, F. (2008): *Argumentation Schemes*. Cambridge University Press, New York.
- Walton, D. & Macagno, F. (2010): "Wrenching from context: The manipulation of commitments", *Argumentation* 24 (3), 283-317.
- Wason, P. (1960): "On the failure to eliminate hypotheses in a conceptual task", *The Quarterly Journal of Experimental Psychology* 12 (3), 129-140.
- (1966): "Reasoning", in Foss, B. M. (ed.), *New horizons in psychology*. Penguin, Harmondsworth, 106-137.
- Woods, J. (2004): *The death of argument: Fallacies in agent based reasoning*. Kluwer Academic Publishers, Dordrecht.

