“I didn’t mean to suggest anything like that!”: Deniability and context reconstruction

Diana Mazzarella

Cognitive Science Centre, University of Neuchâtel

Correspondence

Rue Pierre-à-Mazel 7

2000 Neuchâtel

Switzerland

Email: diana.mazzarella@unine.ch

This is the peer-reviewed version of the following article: Mazzarella, D. (2021), “I didn’t mean to suggest anything like that!”: Deniability and context reconstruction, *Mind & Language*, **https://doi.org/10.1111/mila.12377**

Please quote the published version.

“I didn’t mean to suggest anything like that!”: Deniability and context reconstruction

Verbal communication leaves rooms for interpretative disputes. Speakers can argue about what they mean by their words and negotiate their commitments in conversation. This paper examines the deniability of implicitly communicated contents and addresses the question of what makes an act of denial seem more or less plausible to the addressee. I argue that denials bring about a process of reconstruction of the context of interpretation of the speaker’s utterance and I illustrate how considerations of cognitive utility are the key determinant for distinguishing plausible from merely possible deniability.

KEYWORDS
plausible deniability, implicature, insinuation, context, strategic speaker

1. INTRODUCTION

It is widely assumed that implicit communication enables a risky message to be conveyed while being deniable, disavowable, by the speaker (Fricker, 2012; Pinker, 2007). This feature of implicit communication, often exploited in adversarial conversations or political speeches, has recently attracted the attention of philosophers, psychologists, and linguists. This is because it raises crucial questions about the epistemic value as well as the reputational costs and benefits of all forms of communication that go beyond “stating” or “explicit telling”. For instance, social epistemologists have debated the extent to which deniability represents a challenge for testimonial knowledge, as it appears to undermine the commitments a speaker should undertake to warrant the transmission of beliefs via testimony (Davies, 2019; Peet, 2015). Social psychologists have explored the social consequences of deniability when it comes to trusting speakers as collaborators or advisors (Tenney, Meikle, Hunsaker, Moore, & Anderson, 2019). Linguists have classified defence strategies for commitment denials and enriched the study of deniability with empirical data (Boogaart, Jansen, & van Leeuwen, 2020).

Despite this growing body of work, deniability is still somewhat difficult to pin down. The limits of deniability are blurry, and its workings have not yet received a full-fledged account (but see Camp, 2018, for a significant contribution in this direction). This paper aims at providing a cognitive-pragmatic analysis of deniability and its conditions of possibility by examining the process of utterance reinterpretation which is triggered by an act of denial. In what follows, I first present the notion of plausible deniability introduced by Steven Pinker's seminal work on the strategic advantages of implicit communication (Section 2). This work opens up the issue of what makes deniability *plausible,* as opposed to merely possible. Specifically, it raises the question of when a denial is likely to be perceived as plausible by an audience. The rest of the paper will be devoted to addressing this question. To do this, I proceed as follows: First, I propose to conceive denials as attempts to reconstruct the context of interpretation of the speaker's utterance (Section 3). Second, I illustrate how considerations of cognitive utility concerning the process of context reconstruction can be brought to bear to assess plausible deniability (Section 4). Finally, I discuss how this proposal sheds new light on the relationship between plausible deniability and the strength of communication (Section 5).

2. PLAUSIBLE AND NOT-SO PLAUSIBLE DENIABILITY

The notion of plausible deniability is at the core of Steven Pinker's *theory of the strategic speaker*, which aims to explain the rationale for implicit communication (Lee & Pinker, 2010; Pinker, 2007; Pinker et al., 2008). The theory focuses on so-called "off-record indirect speech acts", which are exemplified by (1) to (4) uttered in appropriate contexts:

(1) Would you like to come up and see my etchings? [a sexual come-on].
(2) Nice house you got there. Would be a real shame if something happened to it [a threat].
(3) We're counting on you to show leadership in our Campaign for the Future [a solicitation of a donation].
(4) Gee, officer, I was wondering whether there might be some way we could take care of the ticket here [a bribe].

(Pinker, 2007, pp. 437–438)

The speaker in (1) to (4) performs a speech act that allows her to implicate some risky content, which is specified in squared brackets, while keeping it "off-record".[1] As "in many implicatures involved in off-record indirect speech acts, the intended message is negative but the literal content is positive or neutral" (Pinker, 2007, p. 439), a strategic speaker may attempt to reduce the potential costs associated with the transmission of the intended message by communicating it indirectly. According to Pinker, indirectness serves this purpose: Whenever the speaker and the addressee have potentially conflicting interests, indirectness leaves it open to the speaker the possibility of denying having had the intention to convey the risky message. This, in turn, allows the speaker to avoid the costs (material or reputational) which she would incur if her message were crafted more explicitly. This point can be illustrated by comparing (4) with an explicit bribe like (5):

---

[1] I am following the convention of referring to the speaker via female pronouns and to the addressee via male pronouns.

(5) If you let me go without a ticket, I'll pay you fifty dollars.

As illustrated in Table 1, when addressed to a corruptible, dishonest officer, the implicated bribe in (4) and the explicit bribe in (5) may both allow the speaker to avoid a traffic ticket. However, if the addressee turns out to be an honest officer (whose interests conflict with the speaker's ones), (5) is a much more costly option than (4).

	Dishonest officer	Honest officer
Don't bribe	Traffic ticket	Traffic ticket
Explicit bribe	Go free	Arrest for bribery
Implicated bribe	Go free	Traffic ticket

Table 1. Payoffs of alternative discourse moves in antagonistic ("honest officer") and non-antagonistic ("dishonest officer") interactions.

Indeed, when challenged by an honest officer, the speaker in (4) can deny having had the intention to bribe the addressee, thus retracting the implicated bribe. The honest officer would lack sufficient evidence to make a bribery charge and prove the speaker's guilt beyond a reasonable doubt. According to Pinker, this possibility, which goes under the name of "plausible deniability", explains the rationale of off-record indirect speech acts.

Plausible deniability relies on the cancellability or defeasibility of implicated meanings (see Lee & Pinker, 2010). As suggested by Grice (1989), conversational implicatures can be cancelled or suspended without logical contradiction:

> [A] putative conversational implicature that *p* is explicitly cancelable if, to the form of words the utterance of which putatively implicates that *p*, it is admissible to add *but not p*, or *I do not mean to imply that p*, and it is contextually cancelable if one can find situations in which the utterance of the form of words would simply not carry the implicature. (Grice, 1989, p. 44)

Consider the following example:

(6) A: Do you wanna come to the cinema tonight?
B: I have to finish writing a paper, but I will come anyway.

While the first part of B's utterance ("I have to finish writing a paper") might invite A to infer that *B does not want to come to the cinema tonight*, this conversational implicature is explicitly cancelled by B's continuation "but I will come anyway".

The relationship between deniability and cancellability/defeasibility is the following: "Though deniability requires defeasibility, the reverse is not true; defeasibility does not require that the speaker have a self-interested reason to cancel the implicature with antagonistic hearers but not with cooperative hearers" (Lee & Pinker, 2010, p. 791). Off-record indirect speech acts are thought of as a particular kind of indirect speech, which is strategically motivated by the speaker's interest to deny (at least part of) her intended meaning if and when the addressee turns out to be antagonistic.

In a similar vein, the philosopher Elizabeth Camp introduces the notion of "insinuation" and distinguishes it from ordinary instances of conversational implicatures based on strategically designed deniability. She suggests that the distinctive feature of insinuations is that if the hearer, an eavesdropper or a third party explicitly attributed to the speaker the intended meaning, the speaker would be prepared and able to deny it coherently (Camp, 2018, p. 45). According to Camp, insinuations involve "the communication of beliefs, requests, and other attitudes 'off record', so that the speaker's main communicative point remains unstated" (Camp, 2018, p. 42). While the notions of "off-record indirect speech act" and "insinuation" are rooted in distinct theoretical frameworks, respectively Speech act theory and Gricean pragmatics, for our discussion, we will ignore these differences and focus on their common reliance on strategically designed deniability.[2]

Given the role attributed to deniability in explaining why speakers engage in off-record indirect speech acts or insinuations, it is worth better understanding how deniability works and what makes deniability plausible. Unfortunately, Pinker's theory of the strategic speaker remains silent with respect to these issues. Pinker (2007) acknowledges that deniability is not always plausible: "The etchings … the leadership, the possibility of accidents, and so on, are transparent ruses, so any 'plausible deniability' is not, in reality, plausible" (Pinker, 2007, p. 453). Most addressees in (1) to (4) would easily read between the lines and understand that the speaker intends to get across a sexual come-on, a solicitation for a donation, a threat or a bribe. Despite this, though, the very possibility of a denial can protect the speaker's interests in an antagonistic interaction. Indirect speech can continue serving its primary strategic function provided that deniability is *possible*, if not plausible. Even when a denial comes across as a bald-faced lie, the addressee is not in a position to call the liar out with absolute certainty, and may well desist doing so to avoid the risks of calling out an innocent speaker (see Pinker 2007, pp. 445–446).

Furthermore, while deniability may not be plausible for the addressee, it could still be plausible for a "virtual audience" (Lee & Pinker, 2010). A third party, who does not participate in the interaction in which the off-record indirect speech act takes place, has no access to a series of extra-linguistic cues that could support the inference to the risky intended content (shared contextual assumptions, tone of voice, facial expression, etc.). From their perspective, the denial may appear as an honest discourse move. When this is the case, the reputation of the strategic speaker within the community will not be at risk.

The distinction between plausible and possible denials evoked by Pinker and colleagues relies on robust intuitions. Indeed, while intuitions concerning the degree of plausibility of a specific denial in one particular context may diverge, it is uncontroversial to assume that denials can be more or less successful, or be perceived as more or less credible. However, Pinker's theory of the strategic speaker does not provide us with any tool to substantiate these robust intuitions with a theoretically sound, and psychologically plausible, analysis of what makes a denial perceived to be more or less plausible from the perspective of the addressee. In what follows, I aim to fill this gap by grounding the distinction between plausible and possible deniability on a cognitive-pragmatic account of utterance interpretation.

---

[2] The assimilation of the two notions appears to be endorsed by Camp herself (2018, p. 42, fn. 4).

3. DENIAL AS CONTEXT RECONSTRUCTION

3.1 Denials

To address the question of the distinction between plausible and possible deniability, I begin by introducing a more precise characterisation of denials. These are discourse moves that consist of denying having had the intention to communicate some implicit content that the addressee inferred and attributed to the speaker as part of her intended meaning. A full-fledged denial typically comprises not only an explicit withdrawal of the target content ("I didn't mean to suggest that *p*") but also the presentation of an alternative interpretation of the utterance ("I only meant that *q*"). Together with the explicit withdrawal (i.e., explicit cancellation), the speaker will be expected to present – or at least be prepared to do so if openly challenged – a situation in which her utterance would not convey the target content (i.e., contextual cancellation).

Through the act of denial, the speaker is thus putting forward a claim of the form "I didn't mean to suggest that *p*, I only meant that *q*", whose likelihood to be accepted as true by the addressee has often important implications for her reputation management. As for any other piece of communicated information, the acceptance of this claim will be modulated by a wide variety of factors. Notably, the addressee's priors about the reliability of the speaker will typically play an important role: The addressee will be more likely to believe the statement of a speaker whose trustworthiness he experienced in the past, or inclined to mistrust a previously inaccurate or deceptive speaker. Furthermore, it is well established that reliability judgements can often be based on general impressions of trustworthiness, delivered by cost-efficient heuristics operating on specific cues, such as age, ethnicity, facial information or accent (for an overview, see Sperber et al., 2010). These variables should certainly be taken into consideration when spelling out a full-fledged psychological account of belief formation via testimony, but I will leave them aside in the discussion of plausible deniability. The reason is that the focus shall be on the question of what makes the denial, *qua* specific discourse move that brings up a reinterpretation of the speaker's utterance ("I only meant that *q*"), seem more or less plausible to the addressee. The attention will be directed to the analysis of the peculiar nature of denials to shed light on the factors that specifically modulate their perceived plausibility, above and beyond those that influence the believability of any piece of communicated information.

With this in mind, it is worth acknowledging that a speaker cannot come up with whatever context would provide a reinterpretation of her utterance to cancel the target implicature. For a denial to be credible, the speaker will need to come up with a context that is a good enough alternative to the actual context of utterance to justify the alleged misunderstanding between the speaker and the addressee.

Camp (2018) makes a similar point. She suggests that "when S is challenged about what she meant, she pretends … to be in a slightly different conversational context C', governed by an alternative set of interpretive assumptions I'" (Camp, 2018, p. 49). The speaker pretends to have addressed her utterance to a hearer who would exploit this set of assumptions to interpret the speaker's utterance and infer her intended meaning. Let us clarify this suggestion with the following example proposed by Camp herself. Imagine that an estate agent addresses a couple of potential buyers from a different racial background or sexual orientation than the local majority with the following question:

(7) Perhaps you would feel more comfortable locating in a more … transitional neighbourhood, like Ashwood?

The speaker in (7) insinuates that the couple should feel uncomfortable in the neighbourhood as they do not conform to the local majority, and should live in a crime-ridden neighbourhood like Ashwood. However, if accused of bigotry, the estate agent could well react with the following denial:

(8) Oh dear me, I didn't mean to suggest anything like that. I only meant that with so many families with young children here, you might not find as many people to socialise with as in a more up-and-coming neighbourhood.

The denial makes it as if the context in which the speaker expected the couple to interpret her utterance was one that included the interpretative assumptions that *childless couples may feel uncomfortable living in neighbourhoods with many families with young children* and that *childless couple may prefer an up-and-coming neighbourhood* (Camp, 2018, p. 49). As these interpretative assumptions *could* have been adopted in interpreting (7), even if the addressees did not employ them, the denial goes through.

Camp's proposal brings us closer to an understanding of the distinction between plausible and possible deniability. Specifically, she maintains that the degree of plausibility of a denial is a function of the degree of epistemic accessibility of the alternative set of interpretative assumptions I' (Camp, 2018, p. 50). In what follows, I develop this suggestion by integrating Camp's philosophical analysis of denials with a cognitive perspective on utterance interpretation. As a first step, the next section examines the notion of context that is brought about by the adoption of this perspective.

3.2 Context and utterance interpretation

The core idea that I will present and develop in what follows is that "determination of the context is not a prerequisite to the comprehension, but a part of it" (Sperber & Wilson, 1982, p. 76). To do this, I will rely on work in cognitive pragmatics by the anthropologist and cognitive scientist Dan Sperber and the linguist Deirdre Wilson, whose collaboration has led to the development of a theory of cognition and communication, known as *relevance theory*, which lays the foundation of much work in contemporary pragmatics.[3]

Comprehension is the process by which the addressee of an utterance infers what the speaker intended to communicate by the use of that utterance. Crucially, the linguistic meaning of the utterance typically underdetermines the speaker's meaning. For instance, a sentence like "It's a sunny day" can be uttered in different occasions to convey different meanings, such as (9a – 9c):

(9) a. A: Do you wanna go to the cinema?
B: It's a sunny day.
b. [Uttered by the accused after being acquitted:]
It's a sunny day.
c. [Uttered under pouring rain to the optimistic friend who organised the pic-nic:]
It's a sunny day.

How do addressees infer the speaker's meaning in all these different circumstances? Building on Grice's intuition that addressees expect the speaker meaning to satisfy some standards of conversational appropriateness (truthfulness, informativeness, etc.), relevance theory suggests that comprehension is guided by expectations of *relevance,* which is defined, in turn, as an optimal balance between the cognitive costs associated with the processing of a stimulus (the utterance) and the cognitive benefits that are achieved as a consequence. All other things being equal, the greater the positive cognitive effects, the greater the relevance; the smaller the processing effort required to derived these effects, the greater the relevance. Positive cognitive effects are changes in an individual's representation of the world that lead to an increase in true representations, more strongly evidenced representations or to a better organisation of information. These effects are the result of the interaction between the incoming information and a "context of available assumptions", which leads to strengthening, revising or eliminating an available assumption or to deriving new contextual implications (Wilson & Sperber, 2004). This raises the question of which available assumptions make up the context of utterance interpretation.

---

[3] For a comprehensive and accessible introduction to the theory, see Clark (2013).

According to Sperber and Wilson, the context of available assumptions that combines with the information carried by the incoming stimulus is *selected* or constructed during the interpretation process. That is, context is not given, but it is chosen by the interpreter, and "context formation is open to choices and revisions throughout the comprehension process" (Sperber & Wilson, 1995, p. 137). But how is the context chosen? Clearly, for each individual at any given time, there is a wide range of assumptions that are in principle available. We thus need an explanation of how the actual context is selected among them. Relevance theory provides an explanation that relies on the very same expectations of relevance that guide the comprehension process as a whole: The actual context is the context that combined with the incoming information satisfies the addressee's expectations of relevance.[4]

As relevance is defined as a function of not only cognitive effects but also cognitive efforts, the context selection should follow a path of least effort (Sperber & Wilson, 1982, p. 76). The addressee starts from the context that is immediately accessible to him—the "initial context"—that comprises the assumptions that are the output of the interpretation of the immediately preceding utterance. The initial context C_0 is the only given context in which the new utterance may be processed. In most cases, though, this initial context provides only a starting point for the construction of the actual context C, which ends up being an expansion of the initial one. An expanded context may include assumptions resulting from the interpretation of previous utterances or previous conversations, encyclopaedic knowledge related to linguistically decoded or inferred concepts, and information from the immediate physical environment. The level of accessibility of these assumptions at any given time will determine the order in which they are accessed. According to Sperber and Wilson, the set of accessible contexts from which C is eventually selected is partially ordered by the inclusion relation, where "order of inclusion corresponds to order of accessibility" (1995, p. 142). The process of expansion, in which every step leads to increased cost in the amount of processing, stops when the context is such that, combined with the new information, it allows the derivation of enough cognitive effects to satisfy the addressee's expectation of relevance. For instance, in (9a) above, the addressee A is likely to expand the initial context C_0 with the assumption that *people prefer to be outdoors when it is sunny* which allows him to derive the contextual implication that *the speaker does not want to go to the cinema*, which satisfies his expectations of relevance by providing an answer to his question.

---

[4] For an in-depth discussion and elaboration of the relevance-theoretic account of context selection, see Assimakopoulos (2017).

The process of context expansion is thus governed by expected relevance: "[I]t is relevance which is treated as a given, and context which is treated as a variable" (Sperber & Wilson, 1995, p. 142). Specifically, relevance theory suggests that the comprehension process is driven by expectations about *optimal relevance*: The addressee is entitled to expect that the utterance would be at least relevant enough to be worth the effort to process it, and the most relevant one compatible with the speaker's means and preferences (Sperber & Wilson, 1995, p. 270). Crucially for our purposes, though, what counts as optimally relevant varies widely across communicative scenarios.

The first source of variation concerns what is taken to be relevant enough in a given circumstance. For instance, this may depend on the degree of intellectual alertness of the addressee: One typically expects more relevance when attending a seminar than when engaging in a casual conversation in a pub (Sperber & Wilson, 1995, p. 161). For this reason, one would be ready to invest more effort into the processing of communicated information in the first than in the second case, and will not be satisfied unless major cognitive effects are achieved. Or, as Gibbs, Kushner, and Mills (1991) showed in the case of metaphor interpretation, interpreters may be influenced by the source of information. Participants in their study were significantly more likely to judge a metaphorical sentence as meaningful, as well as to provide richer metaphorical interpretations of the same sentence, when the target sentences were attributed to famous 20th-century poets than when they were attributed to a computer program that was said to generate them randomly. This finding suggests that interpreters were more likely to invest greater cognitive efforts to make sense of the very same linguistic input when they deemed it more worthy based on information about the source's identity (see also McCallum, Mitchell,

& Scott-Phillips, 2020, for a discussion of the role of credence and institutional setting in the enhanced interpretation of a work of art).

Furthermore, beyond a general expectation of optimal relevance, addressees may have more or less specific expectations about the *way* in which the utterance will achieve relevance when optimally processed, which in turn may depend on the nature of the interaction or the identity of the interlocutor. As Wilson points out, "[i]n many cases (notably in indirect answers to questions, or when a discourse is already underway), he [the addressee] is just as likely to reason backwards from an expected cognitive effect to the context and content that would warrant it" (Wilson, 2004, p. 353). We typically anticipate the way in which our conversational partners will contribute to the conversation (for instance, by settling a question, addressing a doubt, responding to concerns that are shared and at issue). Furthermore, these anticipations may depend on shared assumptions about our partner's interests, preferences, or expertise. For instance, when discussing with your conservative friend about your support to LGBT adoption, her claim "A child needs a family" would be likely to be interpreted as an argument against LGBT adoption (which may prompt you to interpret the utterance as meaning that every child needs a family consisting of two different-sex parents). Your interpretation of the same utterance could be very different if your interlocutor were a friend with progressive political ideas. In this case, you would expect your friend's utterance to achieve relevance by confirming the available assumption that LGBT adoption should be supported. This may lead you to interpret her utterance as meaning that a child's wellbeing is fostered by family caring, including same-sex parenting. That is, in both cases, you may have strong, albeit different, expectations of the specific cognitive effects the utterance is intended to achieve (respectively, contradicting or strengthening the available assumption that LGBT adoption should be supported).

Going back to the process of context construction, it is thus plausible to assume that the context expansions that take place during the interpretative process will be affected by (i) the minimal level of relevance that would satisfy the addressee in a given circumstance and (ii) the existence of specific expectations about the way in which the utterance will achieve optimal relevance in that circumstance. The higher one sets the bar of what counts as relevant enough, the wider the context expansions that are justified by the search for relevance. The more specific the expectations of relevance, the greater their role in leading the search for relevance towards specific cognitive effects.

I have now outlined the process of context construction that is part of the comprehension of any linguistic utterance.[5] The next step is to discuss its implications for the question of the deniability of implicated meanings.

3.3 Context reconstruction

Camp (2018) maintains that denials are attempts to change the context of interpretation of the target utterance from which the addressee inferred a controversial meaning. In line with this, I suggest that denials should be conceived of as *reconstructions* of the actual context of interpretation C. According to Camp (2018, p. 49), denials bring about "a slightly different conversational context C', governed by an alternative set of interpretive assumptions I'". An act of denial is thus an attempt to construct a new context of interpretation C' from the initial context C_0.

Sperber and Wilson conceive of the set of accessible contexts as partially ordered based on their respective accessibility. With this in mind, we can think of the process of context reconstruction as involving a new expansion of the initial context C_0, one which is based on a different partial ordering of the assumptions that are potentially brought to bear in the interpretation process. As a result, the reconstructed context of interpretation will differ from the actual one in that it will either include new contextual assumptions, previously unavailable but made accessible by the act of denial, or exclude some contextual assumptions whose accessibility the speaker will claim not to have foreseen (or both).

With an act of denial, "I didn't mean to suggest that *p*, I only meant that *q*", speakers can propose to reconstruct the context of interpretation by modifying its scope in different ways. First, by broadening it, so that new contextual assumptions can be exploited in deriving relevant contextual implications. This is the case of the following example adapted from Pinker (2007):

(10) A [intending to solicit a donation]: We're counting on you to show leadership in our Campaign for the Future
B: Well, I'm not in a position to make a donation right now.

---

[5] For a discussion of the way in which context selection can be exploited in manipulative discourse containing ill-formed arguments, see Maillat and Oswald (2009) and Maillat (2013). For its implications for presuppositional uses in discourse, see Mazzarella and Domaneschi (2018).

A: Oh, I didn't mean it this way. There are many ways in which you could show your leadership: being a spokesperson for the campaign, volunteering at our campaign offices, recruiting new supporters … Donating is only one of them.

When openly confronted, A can deny having had the intention to solicit a donation from B by broadening the context of interpretation to include contextual assumptions that the addressee had not considered. These may include, for instance, the assumptions *that one can show leadership in a campaign by being a spokesperson, that one can show leadership in a campaign by volunteering at the campaign offices*, etc. A's denial makes these assumptions mutually manifest, and thus sets up a new context of interpretation in which A's original utterance is reinterpreted as an invitation to contribute to the campaign in one of these different ways.

Second, the speaker can manipulate the scope of the context of interpretation by excluding some of the contextual assumptions that the addressee had accessed in constructing it. Once again, let us adapt one of Pinker's (2007) original examples:

(11) A [*intending to invite B to have sex*]: Would you like to come up and see my etchings?
B: I don't hook up on a first date.
A: I'm sorry you took it this way. I didn't mean anything like that. I just wanted to show you my etchings as we talked about them at dinner.

A's denial explicitly eliminates the contextual assumption *offering to come up to one's place is an invitation for sex* that B had inferred in constructing the context of interpretation C. That is, A's denial suggests that the context of interpretation C' should be smaller than B assumed as the only assumptions that the speaker intended to make manifest involved A's etchings and A's willingness to show them to B.

Finally, a denial can invite a reconstruction of the context of interpretation which both exploits new contextual assumptions in the derivation of the intended cognitive effects and excludes some of the assumptions which were part of the actual context of interpretation: That is, the speaker invites the addressee to replace the latter with the former. Consider this modified version of Pinker's (2007) bribe example:

(12) A [*intending to bribe the officer*]: Gee, officer, I was wondering whether there might be some way we could take care of the ticket here.
B: You know I can arrest you for a bribe.

> A: Oh, I didn't mean it that way, officer. I was just wondering whether we could use a mobile terminal to pay the fine.

In this case, A's denial pushes B back to the initial context C_0 and invites him to access a new set of assumptions concerning payments via mobile terminals, which had not been previously contemplated and whose implications directly contradict B's previous conclusions about A offering him a bribe.

To sum up, denials bring about a new context of interpretation C', through a process of context reconstruction from the initial context C_0. This process is based on a rearrangement of the available contextual assumptions, in line with what the speaker pretends to have had in mind when producing her utterance. More technically, this amounts to making the case for a different partial ordering of the set of potential contexts available for interpretation.

4. PLAUSIBLE AND IMPLAUSIBLE CONTEXT RECONSTRUCTIONS

Building on the notion of context reconstruction, in this section, I consider the question of when the reinterpretation of the speaker's utterance, which is forced by the act of denial, is perceived as plausible by the audience. The proposal I will develop is grounded on the idea that its perceived plausibility crucially depends on the cognitive utility associated with the intended context reconstruction.

Hinting to considerations of cognitive utility, Camp (2018) claims that the degree of plausible deniability is a function of the degree of epistemic accessibility of the alternative set of interpretative assumptions I' (Camp, 2018, p. 50). However, the notion of epistemic accessibility is left underspecified. In what follows, I develop Camp's claim by suggesting that the plausibility of a denial is modulated by considerations of both cognitive effects (e.g., licensing of implicated conclusions) and cognitive efforts (e.g., accessibility/salience of contextual assumptions). The aim is to provide a way to anchor the distinction between plausible and merely possible deniability to the nature of the cognitive process of utterance interpretation.

Let us focus on cognitive effects first. Consider the following example from Grice (1989, also discussed by Camp, 2018) of a letter of recommendation for a philosophy job:

(13) Mr X’s command of English is excellent, and his attendance at tutorials has been regular.

As the explicit content of the letter does not satisfy the expectations of relevance of a selection committee, the addressee is invited to infer that the author meant to suggest, or insinuate, that Mr X is not a good philosopher. It is the derivation of this implicature that makes (13) relevant enough to be worth considering. Imagine that the author, openly confronted by a member of the committee, denies having had the intention to communicate a negative judgement about the candidate. For instance, the author insists that the only intended meaning corresponds to what has been explicitly asserted. This is an example of a denial that attempts to narrow down the scope of the context of interpretation by excluding those assumptions licensing the derivation of the risky implicature. However, in this case, the reconstructed context C’ does not warrant the derivation of enough cognitive effects to make the utterance worth processing. By restricting the context of interpretation in such a way that no assumptions concerning the marginality of the qualities mentioned in (13) for the evaluation of a philosopher could be used to derive the implicature at issue (*Mr X is a bad philosopher*), the author would fail to meet any standard of conversational relevance. While the denial is possible, it would not be plausible. Plausible deniability requires that the reinterpretation of the utterance, which involves the reconstructed context C’, should be relevant to the addressee (as opposed to irrelevant) and achieve some cognitive effects.

This point is acknowledged by Haugh (2013), who suggests that the degree of plausibility of a denial is a function of the degree of saliency of the implicature, as well as its status as primary or secondary meaning: “In cases where something has clearly been implied (i.e. implicature as a primary meaning), the speaker can usually only clarify what was implied. In cases where it is open to interpretation whether something has been implied (i.e. implicature as secondary meaning), then implicatures may be more readily denied or retracted as well” (Haugh 2013, p. 147). In other terms, while the speaker can deny having had the intention to communicate a certain implicature (*that Mr X is a bad philosopher*), if the relevance of the utterance depends on the derivation of *some* implicature (offering a relevant judgement about Mr X as a philosopher), a denial putting forth a context C’ which does not allow the derivation of any relevant implicature will be perceived as hopelessly implausible.

A plausible denial thus requires putting forth a reconstructed context which warrants the achievement of at least some cognitive effects. This achievement is meant to justify the cognitive effort invested by the addressee into the processing of the speaker’s utterance. For this reason, the speaker cannot suggest a blatantly irrelevant re-interpretation and expect the addressee to believe that she had the intention to communicate it.

Suppose, though, that the statement in (13) is uttered under different circumstances. For instance, as part of an informal conversation between the speaker and another philosophy professor, who is a member of the selection committee in charge of evaluating Mr X. As before, suppose that, when openly confronted about her intention to express a negative evaluation of Mr X as a philosopher, the speaker denies having insinuated any negative judgement and exclusively commits to the truth of the asserted content. Is the speaker's denial more likely to be perceived as credible under these circumstances than in the reference letter scenario? Indeed, this seems to be the case. To understand why, it is worth focusing on the observation, discussed in Section 3.2 above, that what counts as an optimally relevant interpretation depends on (i) the minimal level of relevance that the addressee expects the utterance to achieve, and (ii) the presence or absence of specific expectations about how the utterance will achieve cognitive effects. Concerning this, Sperber and Wilson suggest that "[o]n various social occasions, the expected level of relevance is culturally defined. In the course of a conversation, the level can be adjusted, increased or decreased, one step at the time. The addressee may make manifest the minimal level of relevance he expects: by asking a question, for instance" (Sperber & Wilson, 1995, p. 161). The reference letter scenario is one in which the social conventions related to the practice of acting as a referee make it mutually manifest to the speaker and the addressee that the receiver of the letter is expecting the reference to provide an answer to the open question of whether the applicant is a suitable candidate for the job and that anything else would not be considered as relevant enough. In contrast with this, the informal conversation scenario seems to leave open the possibility—invoked by the denial—that the speaker could have *attempted* to be optimally relevant, despite failing to do so.[6] First, the speaker may argue, the nature of the exchange may have led him to think that only a modicum of relevance would have been enough (e.g., a modicum achieved by confirming the assumption that Mr X is well-integrated into the regular activities of the department). Second, if the addressee had not explicitly asked for an evaluation of Mr X's qualities as a philosopher, the speaker may insist to have failed to recognize the specificity of the interlocutor's expectations about the cognitive effects of her utterance.

---

[6] For a discussion of the notion of "attempted optimal relevance" as opposed to "actual optimal relevance", see Wilson (2005, p.143).

With this in mind, we can thus maintain that a plausible denial will guarantee some cognitive utility in terms of the achievement of enough cognitive effects to offset the amount of effort invested by the addressee in processing the speaker's utterance. However, as what counts as relevant enough can vary dramatically across conversational contexts, the plausibility of the denial can be preserved if the speaker can reasonably argue to have misjudged her interlocutor's expectations. This leads to some specific empirical predictions. For instance, that a denial will be perceived as more plausible in conversational settings in which there are no culturally defined expectations about the minimal level of relevance an utterance is expected to achieve, or that a denial would be perceived as less plausible when the interlocutor had attempted to make manifest this minimal level of relevance in the course of the conversation (e.g., by asking an explicit question).

Let us now move to considerations of cognitive efforts and explore the way in which they affect the degree of perceived plausibility of a denial. To being with, it is crucial to emphasise that "the amount of processing tends to remain roughly constant throughout a stretch of discourse" (Sperber & Wilson, 1982, p. 77). For instance, when engaging in casual chitchat, one does not expect to invest significant effort in interpreting the interlocutor's utterances. For this reason, in informal conversations, interlocutors will not typically scrutinise all the implications of each other's statements, as this would be unjustified given the nature of the conversation. In contrast, when reading a scholarly work or interpreting a piece of literature, the expectation that greater attention will lead to an increase in relevance justifies in-depth processing of each piece of text and a constant expansion of the context of interpretation (Furlog, 1995). The stability in the amount of processing that the addressee is likely to invest within the same stretch of discourse predicts and explains instances of underprocessing or overprocessing: For example, a veiled reference to a difficult personal situation may be missed in the context of chitchat because of underprocessing, but would not go unnoticed in a therapy session. Vice versa, a banal observation from a therapist might be overprocessed in vain by a client eager to find it more meaningful than it was meant to be. This lack of flexibility in the amount of processing has significant implications on the process of context construction (and reconstruction). As a result, "the search for an adequate context tends to remain within *a predictable and generally narrow domain*. In trying to maximise relevance, the speaker must adapt to this fact, and the hearer can assume that he has" (Sperber & Wilson, 1982, p. 77; *my emphasis*).

Building on this, I suggest that the lack of flexibility in the amount of processing throughout the same conversation, together with its foreseeability, constrain the class of alternative contexts that speakers can plausibly put forward as part of their denials. In other terms, a reconstructed context C' will appear as a plausible context of reinterpretation if and only if it falls within a sufficiently "narrow domain", of which the actual context of interpretation is part and which could have been reasonably accessed by the addressee. That is, the amount of processing that the speaker might have expected the addressee to invest to access C' should be only marginally smaller or greater than the one which underpinned the construction of the actual context of interpretation C.

To illustrate this point, let us go back to examples (7) and (8), uttered by a bigoted estate agent to a childless couple from a different racial background than the local majority (reported below for convenience):

(7) Perhaps you would feel more comfortable locating in a more … transitional neighbourhood, like Ashwood?

(8) Oh dear me, I didn't mean to suggest anything like that. I only meant that with so many families with young children here, you might not find as many people to socialise with as in a more up-and-coming neighbourhood.

Let us consider an alternative denial, such as:

(14) Oh dear me, I didn't mean to suggest anything like that. I only meant that with so many stores selling only primary necessities, you might not find as many little shops and boutiques to make your purchases as in a more up-and-coming neighbourhood.

Denials (8) and (14) involve the reconstruction of two different contexts, which are the result of alternative expansions. Crucially, the first consists of an expansion from the initial context C_0 that adds the assumption that *childless couples may feel uncomfortable living in neighbourhoods with many families with young children*, while the latter requires the addition of the assumption that *couples may dislike residing in neighbourhoods with limited shopping options for non-essential items*. The plausibility of the two denials depends on the accessibility of these assumptions starting from the initial context C_0: The more accessible the relevant interpretative assumption, the more plausible the denial. If we imagine that assumptions about the demographics of neighbourhoods are more salient to the addressee than assumptions about the availability of non-essential shops and boutiques, it follows that the addressee will perceive denial (8) as more plausible than denial (14). Finally, it is worth pointing out that the perceived plausibility of the denial in (8) will also depend on the salience of the assumption it invokes (e.g., *childless couples may feel uncomfortable living in neighbourhoods with many families with young children*) relative to the salience of the assumption that contributed to the original interpretation of the speaker's utterance in (7) (e.g., *Black couples may feel uncomfortable to live in neighbourhoods with many white families*). The greater the relative salience of the latter over the former, the greater the amount of cognitive effort which would have been required for accessing them and, as a result, the less plausible the denial.

By her act of denial, the speaker modifies the accessibility of a potential context of interpretation, by making mutually manifest (or more manifest) contextual assumptions that the addressee had not mobilized in his first interpretation of the utterance. While a denial can in principle raise the manifestness of any contextual assumption, the perceived plausibility of a denial depends on the degree of accessibility of these assumptions in the cognitive environment of the addressee at the time of the initial interpretation. Given that the addressee is entitled to expect the speaker to be optimally relevant, and that relevance is a negative function of cognitive effort, the alternative interpretation put forth by the denial as the intended one could not have required an amount of processing radically different from the one the addressee devoted to the interpretative process (e.g., it cannot rely on highly unavailable assumptions). For this reason, the proposal is that a plausible denial should involve the construction of a new context C', whose assumptions would have required to the addressee only a marginally different (smaller or greater) amount of cognitive effort than the one invested for the construction of the actual context of interpretation C. The need for greater flexibility in the amount of processing invested could only be reasonably justified under special circumstances. For instance, while the speaker may be unlikely to misjudge the degree of attention the addressee will pay to her utterances over a stretch of discourse, mistakes about the degree of intellectual alertness of the interlocutor may be more likely to occur (and to be condoned) at the beginning of a conversation, especially when the nature or purpose of the conversational exchange is not yet clear. For this reason, unless the speaker may convincingly argue for such a misjudgement, a plausible denial will typically rely on a reconstructed context that falls within a predictable and generally narrow domain.

The examples discussed above show that judgements about the plausibility of a denial will depend on consideration of both the cognitive effects achieved by, and the cognitive efforts potentially involved in, the reconstruction of the context of interpretation C'. With this in mind, we can thus conclude that the distinction between plausible and merely possible denials is better thought of in terms of a continuum of cases, which ranges from plausible context reconstructions characterised by higher cognitive utility (enough cognitive effects and relatively stable cognitive costs) to implausible context-reconstructions that warrant irrelevant reinterpretations of the speaker’s utterance (merely possible denials). As Camp clearly states, the speaker of a denial is “pretending to address *U* to a possible hearer H_P who would sincerely employ the alternative assumptions *I*’” (Camp, 2018, p. 50) to derive the speaker’s intended meaning. Considerations of cognitive utility provide us with a measure of the magnitude of this pretence and thus offer us a framework to understand when the pretence is blatant and when it has better chances to go unnoticed. If the suggested reinterpretation of the utterance falls short of conversational relevance only marginally, the alleged misunderstanding between the speaker and the addressee appears more justifiable. Indeed, while the speaker might have been mistaken in her “nuanced sensitivity to interpretative salience and relevance” (*Ibid.*), she was not completely off.

To conclude, it is worth pointing out that the perceived plausibility of the denial is not simply a function of the cognitive utility of the re-interpretation of the speaker’s utterance, but can also depend on the identity of the interlocutors and the relationship between them. Indeed, the extent to which a speaker might have been reasonably mistaken in predicting the addressee's interpretation of her utterance is constrained by the relationship between the interlocutors and their shared background knowledge. This is because predictions about how one’s statement will achieve relevance for an addressee often depend not only on circumstantial factors (e.g., the question under discussion) but also on our prior knowledge about the interlocutor’s background, areas of expertise, interests, and so forth. I illustrate this point with another example adapted from Grice’s (1989):

(15) A: Shall we invite Steve to join us for dinner?
B: Steve is meeting a woman tonight.

Grice (1975/1989, p. 37) observes that an utterance of (15) “would normally implicate that the person to be met was someone other than X’s wife, mother, sister, or perhaps even close platonic friend”). Imagine, though, that the speaker B denies having had the intention to communicate this implicature, for instance, as in (16):

(16) I didn’t mean it that way! I just wanted to point out that Steve already had some plans for the night.

Compare the plausibility of the denial in the following two conversational settings. In the first one, the addressee A is Gina, who is notoriously in love with Steve. In the second one, the addressee A is B’s flatmate, who has met Steve once or twice. Could B have failed to predict that the addressee would infer that Steve had a date that night? Hardly, if the addressee was Gina; possibly, if the addressee was someone who had no particular interest in Steve’s romantic life.

To sum up, the degree of plausibility of a denial depends on the extent to which the reconstructed context of interpretation put forth by the speaker represents a plausible alternative to the actual context of interpretation. An alternative is plausible if it still warrants the derivation of enough cognitive effects subject to the investment of reasonable cognitive efforts. As discussed above, the amount of effort that the speaker could have reasonably expected the addressee to invest in processing her utterance is often a function of the conversational setting, and the stability in the expected amount of processing over a stretch of discourse further constrains this expectation. Furthermore, as shown by example (16), the fact that interlocutors share relevant background knowledge should in principle facilitate the recognition of the most accessible contextual assumptions that will be exploited by the addressee in interpretation, and thus represents another constraint for the credibility of a denial.

5. PLAUSIBLE DENIABILITY AND THE STRENGTH OF COMMUNICATION

In this section, I aim at discussing the relationship between plausible deniability, on the one hand, and the strength of communication, on the other hand. I intend to show that the analysis of the cognitive processes of context construction and reconstruction provides us with a way to understand why weakly communicated contents are typically assumed to be more deniable (see Sternau et al., 2017 for empirical evidence).

Communication is not a yes/no matter—that is, either a specific thought is communicated or it is not—but can be conceived as a matter of degree (Wilson, 2011). Pragmaticists often distinguish between so-called “strong” and “weak” implicatures (Wilson & Sperber, 2004) based on their role in the comprehension process and their relationship with the speaker’s intentions. A strong implicature is a proposition that is strongly implicated by the speaker, that is, one that the speaker has in mind and intends the addressee to derive. The recovery of a strong implicature is indeed essential to arrive at a relevant interpretation of the speaker’s utterance. In contrast with this, an implicature is only weakly implicated (that is, it is a weak implicature) if its derivation might—but need not to—contribute to the recovery of an interpretation that satisfies the addressee's expectations of relevance. Typically, a weak implicature is part of a range of similar possible implicatures, any one of which would do. In line with this, a weak implicature is not necessarily foreseen by the speaker (the speaker need not entertain *that specific* implication, or have the intention that the addressee derives it). As Sperber and Wilson (1995) suggest:

> [T]here is a continuum of cases, from implicatures which the hearer was specifically intended to recover to implicatures which were merely intended to be made manifest, and to further modifications of the mutual cognitive environment of speaker and hearer that the speaker only intended in the sense that she intended her utterance to be relevant, and hence to have rich, and not entirely foreseeable, cognitive effects. (Sperber & Wilson, 1995, p. 201)

For our discussion, it is worth noting that the strength of the implicature is claimed to affect the degree of the hearer's *responsibility* in deriving it. Specifically, the weaker the implicature, the greater the hearer's responsibility for the interpretative choices he makes. Conversely, the stronger the implicature, the more responsibility falls on the speaker (Sperber & Wilson, 2006). Table 2 summarises the essential features that distinguish strong from weak implicatures.

Strong implicature	**Weak implicature**
Essential for comprehension	Optional for comprehension
Typically unique	Typically part of a range
Foreseeable	Not necessarily foreseeable
Speaker's responsibility	Hearer's responsibility

Table 2. Implicature strength

Recent experimental work from Sternau et al. (2017) has demonstrated that weak implicatures are considered to be more easily deniable than strong implicatures. That is, a speaker can more easily deny having had the intention to communicate a weak implicature than a strong one.

Based on the analysis of plausible deniability offered in this paper, I suggest that we can think of the strength of an implicature as determined by the extent to which the process of context construction is constrained given the speaker's utterance and the addressee's expectations of relevance. If the relevance of the speaker's utterance crucially depends on the construction of a specific context, one which only would licence the derivation of enough relevant conclusions, these conclusions are strongly implicated. On the other hand, if a whole range of alternative contexts C_1-C_n can in principle lead to the derivation of conclusions that would be relevant to a comparable extent, when one of these conclusions is derived, it will be weakly implicated.

With this in mind, it is easy to see why the degree of plausible deniability of an implicature is inversely proportional to its strength. In the case of weak implicature, the speaker can more easily reconstruct the context of interpretation by putting forth any other alternative context belonging to C_1-C_n. This would likely result in a relevant (though not optimal) interpretation of the utterance, thus making the denial appear relatively plausible.

Interestingly, the relationship between plausible deniability and strength of communication sheds light on the possibility of extending the account presented in this paper to other weak, albeit not implicit, forms of communication. Post-Gricean approaches to pragmatics, like relevance theory, have long acknowledged the crucial contribution of pragmatic inference to the recovery of what is *explicitly* (and not only implicitly) communicated by a speaker (see, e.g., Carston, 2002). To illustrate this, consider the following example from Wilson and Sperber (2002):

(17) a. Alan: Do you want to join us for supper?
b. Lisa: No thanks. I've eaten.
c. Lisa: No, thanks. I've already eaten supper.
d. Lisa: No, thanks. I've already eaten tonight.
e. Lisa: No, thanks. I've already eaten supper tonight.

In this exchange, any version of Lisa's utterance would be taken to explicitly communicate that Lisa has already eaten supper that night and to implicate a negative answer to Alan's question. Crucially, though, these alternative utterances vary in the extent to which they involve pragmatic inference beyond linguistic decoding. The contribution of pragmatic inference decreases from (17b) to (17e): Like for implicatures, "[e]xplicatures can be weaker or stronger, depending on the degree of indeterminacy introduced by the inferential aspect of comprehension" (Wilson & Sperber, 2002, p. 249). This point has implications with respect to the degree of responsibility of the addressee's in deriving the speaker meaning: the weaker the explicature, the greater his responsibility. Once again, then, strength appears to be linked with the possibility of a denial. Following Wilson and Sperber's discussion of this example, suppose Alan replies to Lisa:

(18) What you just said is false: I happen to know that you haven't eaten a thing tonight!

While Lisa could not deny having had the intention to communicate that she had supper that night had she uttered (17e), a denial seems to be at least possible for an utterance communicating a weaker explicature like (17b). Furthermore, and in line with the account of plausible deniability developed in this paper, the plausibility of the denial may depend on the relative accessibility of the new contextual assumptions that are brought to bear in the reinterpretation process. For instance, Lisa's explanation that she had such a nutritious snack that afternoon that she did not feel like eating anything else may come across as more plausible than the alternative explanation that since she had such a heavy dinner the night before she would rather fast that day. As both assumptions warrant the derivation of the intended implicature, but the former is potentially more accessible than the latter, considerations of cognitive utility will weigh in its favour.

While this analysis of plausible deniability can in principle be applied to a broader range of pragmatic phenomena, off-record indirect speech acts and insinuations represent the prototypical types of strategic communication employed in adversarial or non-cooperative circumstances as their implicit nature has been shown to reduce the speaker's accountability towards the message communicated (Mazzarella et al., 2018).

6. CONCLUSIONS

Communication often takes place in adversarial or non-cooperative circumstances, in which the interests of the interlocutors do not align with one another. In these circumstances, strategic speakers may communicate risky messages while attempting to preserve deniability. This requires a readiness to retract the message while providing an alternative, plausible, interpretation of their linguistic utterances. This paper aimed at elucidating the notion of plausible deniability. The analysis presented here is grounded in a cognitively-oriented understanding of the inferential nature of communication, one which emphasises the active process of context construction in the interpretation of any linguistic utterance. Based on this understanding, I suggested that denials should be conceived of as attempts to reconstruct the context of interpretation and highlighted how considerations of cognitive utility can orient this reconstruction and determine its perceived plausibility.

The notion of plausible deniability has been discussed by philosophers, linguists and psychologists to understand the strategic advantages of implicit communication and their implications for the justification of testimonial knowledge. However, the question of what makes a denial plausible had not yet received a compelling answer. The primary goal of this paper was to provide such an answer. A better understanding of plausible deniability can contribute to this interdisciplinary body of work by offering a cognitive framework to philosophical analysis as well as by anchoring empirical investigation to theoretically motivated premises.

ACKNOWLEDGEMENTS

I am very grateful to the editor and the two anonymous reviewers for their insightful comments and suggestions.

REFERENCES

Assimakopoulos, S. (2017). Context in relevance theory. In Blochowiak, J., Grisot, C., Durrleman, S., & Laenzlinger, C., (Eds.), *Formal models in the study of language: Applications in interdisciplinary contexts* (pp. 221–242). Cham: Springer.

Boogaart, R., Jansen, H., & van Leeuwen, M. (2020). “Those are your words, not mine!” Defence strategies for denying speaker commitment. *Argumentation*. https://doi.org/10.1007/s10503-020-09521-3

Camp, E. (2018). Insinuation, common ground. In Fogal, D., Harris D. W., & Moss, M., (Eds.), *New work on speech acts* (pp. 40–65). New York, NY: Oxford University Press.

Cartson, R. (2002). *Thoughts and utterances*. Oxford: Blackwell Publishers.

Clark, B. (2013). *Relevance theory*. Cambridge: Cambridge University Press.

Davies, A. (2019). Testimony, recovery and plausible deniability: A response to Peet. *Episteme, 16*(1), 18–38. https://doi.org/10.1017/epi.2017.15

Fricker, E. (2012). Stating and insinuating. *Proceedings of the Aristotelian Society Supplementary Volume LXXXVI,* 61–94. https://doi.org/10.1111/j.1467-8349.2012.00208.x

Furlog, A. (1995). *Relevance theory and literary interpretation* (Unpublished doctoral dissertation), UCL, London, United Kingdom.

Gibbs, R. W., Kushner, J. M., & Mills, W. R. (1991). Authorial intentions and metaphor comprehension. *Journal of Psycholinguistic Research, 20*(1), 11–30. https://doi.org/10.1007/BF01076917

Grice, H. P. (1989). *Studies in the way of words*. Cambridge, MA: Harvard University Press.

Haugh, M. (2013). Implicature, inference and cancellability. In Capone, A., Lo Piparo, F., & Carapezza, M., (Eds.) *Perspectives on pragmatics and philosophy* (pp. 133-151). Cham: Springer.

Lee, J. J., & Pinker, S. (2010). Rationales for indirect speech: The theory of the strategic speaker. *Psychological Review, 117*(3), 785–807. https://doi.org/10.1037/a0019688

Maillat, D., & Oswald, S. (2009). Defining manipulative discourse: The pragmatics of cognitive illusions. *International Review of Pragmatics, 1*(2), 348–370. https://doi.org/10.1163/187730909X12535267111651

Maillat, D. (2013). Constraining context selection: On the pragmatic inevitability of manipulation. *Journal of Pragmatics, 59,* 190–199. https://doi.org/10.1016/j.pragma.2013.07.009

Mazzarella, D., & Domaneschi, F. (2018). Presuppositional effects and ostensive-inferential communication. *Journal of Pragmatics, 138,* 17–29. https://doi.org/10.1016/j.pragma.2018.09.012

Mazzarella, D., Reinecke, R., Noveck, I., & Mercier, H. (2018). Saying, presupposing and implicating: How pragmatics modulates commitment. *Journal of Pragmatics, 133,* 15–27. https://doi.org/10.1016/j.pragma.2018.05.009

McCallum, K., Mitchell, S., & Scott-Phillips, T. (2020). The art experience. *Review of Philosophy and Psychology, 11*(1), 21–35. https://doi.org/10.1007/s13164-019-00443-y

Peet, A. (2015). Testimony, pragmatics, and plausible deniability. *Episteme, 12*(1), 29–51. https://doi.org/10.1017/epi.2014.31

Penco, C., & Domaneschi, F. (2013). *What is said and what is not: The semantics/pragmatics interface*. Stanford: CSLI Publications.

Pinker, S. (2007). The evolutionary social psychology of off-record indirect speech acts. *Intercultural Pragmatics, 4*(4), 437–461. https://doi.org/10.1515/IP.2007.023

Pinker, S., Nowak, M. A., & Lee, J. J. (2008). The logic of indirect speech. *Proceedings of the National Academy of Sciences, 105*(3), 833–838. https://doi.org/10.1073/pnas.0707192105

Sperber, D., Clément, F., Heintz, C., Mascaro, O., Mercier, H., Origgi, G., & Wilson, D. (2010). Epistemic vigilance. *Mind & Language, 25*(4), 359–393. https://doi.org/10.1111/j.1468-0017.2010.01394.x

Sperber, D., & Wilson, D. (1982). Mutual knowledge and relevance in theories of comprehension. In Smith, N., (Ed.), *Mutual Knowledge* (pp. 61–87). London: Academic Press.

Sperber, D., & Wilson, D. (1995). *Relevance: Communication and cognition*. Oxford: Blackwell Publishing.

Sperber, D., & Wilson, D. (2006). Pragmatics. In Jackson, F., & Smith, M., (Eds.), *Oxford Handbook of philosophy of language* (pp. 468-501). Oxford: Oxford University Press.

Sternau, M., Ariel, M., Giora, R., & Fein, O. (2017). Deniability and explicatures. In Giora, R., & Haugh, M., (Eds.), *Doing pragmatics interculturally* (pp. 97–120). Berlin/Boston: Walter de Gruyter GmbH.

Tenney, E. R., Meikle, N. L., Hunsaker, D., Moore, D. A., & Anderson, C. (2019). Is overconfidence a social liability? The effect of verbal versus nonverbal expressions of confidence. *Journal of personality and social psychology*, *116*(3), 396–415. https://doi.org/10.1037/pspi0000150

Wilson, D. (2004). Relevance and lexical pragmatics. *UCL Working papers in Linguistics, 16,* 343–360.

Wilson, D. (2011). Relevance and the interpretation of literary works. In Yoschimura, (Ed.), *Observing linguistic phenomena: A festschrift for Seiji Uchida* (3-19). Tokyo: Eihohsha.

Wilson, D., & Sperber, D. (2002). Truthfulness and relevance. *Mind*, *111*(443), 583-632. https://doi.org/10.1093/mind/111.443.583

Wilson, D., & Sperber, D. (2004). Relevance Theory. In Horn, L.H. & Ward, G.L. (Eds.), *Handbook of Pragmatics* (pp. 607-632). Oxford: Blackwell.