

Hearing Meanings: The Revenge of Context

Luca Gasparri and Michael Murez | August 30, 2019

(penultimate draft, final version published in *Synthese*)

Abstract

According to the perceptual view of language comprehension, listeners typically recover high-level linguistic properties such as utterance meaning without inferential work. The perceptual view is subject to the Objection from Context: since utterance meaning is massively context-sensitive, and context-sensitivity requires cognitive inference, the perceptual view is false. In recent work, Berit Brogaard provides a challenging reply to this objection. She argues that in language comprehension context-sensitivity is typically exercised not through inferences, but rather through top-down perceptual modulations or perceptual learning. This paper provides a complete formulation of the Objection from Context and evaluates Brogaard's reply to it. Drawing on conceptual considerations and empirical examples, we argue that the exercise of context-sensitivity in language comprehension does, in fact, typically involve inference.

1. Introduction

John is a native speaker of English and has never heard a word of French in his life. Leo is a native speaker of French. Suppose that John and Leo are presented with the same utterance of the French sentence in (1).

- 1) Marie est allée à l'école il y a deux heures.

John and Leo will hear the same speech sounds. However, knowing the language, Leo will be able to differentiate the speech stream into segments, recognize what words have been uttered, reconstruct their grammatical organization, and recover the meaning of the sentence: Marie went to school two hours ago. John, by contrast, lacks the background linguistic knowledge required to perform these operations. As a result, his experience of the utterance will be different: John will apprehend the speech sounds and their audible properties, but not their meaning.

This example suffices to illustrate a simple and uncontroversial fact: the experience of perceiving utterances in a known language and the experience of perceiving utterances in an unknown language are remarkably different (see, a.o., Tye 2000; Prinz 2006). Among the differences involved, one concerns in particular the appearance of high-level semantic properties. When we experience utterances in a known language, our experience features the appearance of high-level semantic properties such as

utterance meaning. When we experience utterances in a language we do not know, by contrast, our experience does not feature the appearance of such properties. Call this the Semantic Contrast.

Issues surrounding the Semantic Contrast and the nature of the phenomenology involved in the apprehension of semantic properties have received considerable attention recently. Over the years, three main rival views have emerged: Conservatism, Moderatism, and Liberalism.

According to Conservatism, competent listeners have a perceptual appearance only of the low-level properties of utterances, such as duration, amplitude, and pitch (e.g., Strawson 1994). On this view, the Semantic Contrast is a difference in cognitive representation which does not impact perceptual phenomenology. According to Moderatism, competent listeners have a perceptual appearance of the low-level and mid-level, language-specific features of utterances, such as their phonemic and syllabic structure (e.g., O’Callaghan 2011; Reiland 2015). Proponents of Moderatism disagree with proponents of Conservatism on the boundaries of perceptual phenomenology, which they take to include the appearance of mid-level properties, but agree that the Semantic Contrast is a difference in cognitive representation which does not impact perceptual phenomenology. According to Liberalism, finally, competent listeners have a perceptual appearance of the low-level, mid-level, and high-level properties of utterances, such as their meaning (e.g., Siegel 2006; Bayne 2009). On this view, in contradistinction to Conservatism and Moderatism, the Semantic Contrast is a difference in perceptual phenomenology.

Let us focus on Liberalism. As stated, Liberalism is a thesis about the *outcome* of meaning recovery. It claims that the processes underlying our interpretation of utterances in everyday language comprehension terminate with a state which is best characterized as a perceptual appearance of meaning. As such, Liberalism can be conjoined without obvious contradiction with various different accounts of the *process* of meaning recovery itself, viz. the sort of transitions between prior states which culminate in our comprehension of utterances. However, some of its proponents commit to a specific account of linguistic processing, which, they argue, is particularly attractive when combined with their picture of the phenomenology of language: the view according to which the process of meaning recovery is, typically, thoroughly non-inferential (henceforth, “Non-Inferentialism”) (Brogaard 2018; 2019). Our target in this paper will be precisely the conjunction of the Liberal claim that the outcome of the process of meaning recovery is best characterized as a perceptual appearance of meaning, with Non-Inferentialism about the process of meaning recovery. This conjunction (Liberalism + Non-Inferentialism) we will refer to as the Perceptual View.

The Perceptual View is unorthodox, and somewhat counterintuitive if measured against some widely shared assumptions about the nature of language comprehension. For example, the view appears to go against the received view in linguistics that we hear utterances and recover their content by making tacit inferences based on our knowledge of the language (what linguists in the generative tradition characterize as the computational labor of our internal grammar; see, e.g. Ludlow 2011; Isaac and Reiss 2013). Tasks such as recovering the syntactic tree of a sentence and assembling the semantic values of its lexical nodes via compositional processes seem to rely on operations that few would feel comfortable characterizing as perceptual.

Nevertheless, the Perceptual View does come with some interesting perks. First, it has the advantage of providing an attractively simple and cohesive picture of the relationship between linguistic processing and linguistic phenomenology, which are both ultimately grounded in perception. Furthermore, by ascribing the apprehension of semantic properties to perception, the account produces a compelling explanation of the subjective feeling of immediacy and fluency we experience when we grasp the meaning of utterances in a known language. The Perceptual View also appears to avoid a familiar epistemological worry affecting theories committed to Inferentialism about meaning recovery: namely, that we cannot really know the meaning of the utterances we hear, because the available evidence about speakers' linguistic behavior underdetermines facts about the meaning of their utterances (Davidson 1973). According to the Perceptual View, by contrast, when we hear utterances in a familiar language, the experience confers immediate, basic justification on our beliefs about what was said, and this is so because we have a non-inferential capacity to be perceptually acquainted with the meaning of such utterances (Brogaard 2017; 2018).¹

So there are good reasons to seriously consider the Perceptual View. However, in order to argue that considerations of this sort actually tip the balance in its favor, one would first need to show that the account doesn't suffer from any more fundamental issues. And yet, there is a major worry lurking in the background: context-sensitivity.²

¹ Notice that this advantage rests on the prior assumption that the capacity to "immediately justify" exhibited by appearances of utterance meaning is best accounted for within a perceptual model of utterance comprehension. Recent work by Balcerack Jackson (2019) has argued that this assumption can be resisted: even non-perceptual models of utterance comprehension can account for the fact that the experience as of a speaker asserting that *p* normally confers immediate justification for believing that the speaker has asserted that *p*. In what follows, we won't dwell on this issue, though we briefly return to epistemological matters in discussing the notion of inference. For more on the epistemology of language comprehension, see, e.g., Longworth (2008).

It is uncontroversial that everyday language comprehension relies pervasively on the appreciation of contextual facts. For example, consider (2).

- 2) The burglar threatened the student with a knife.

This sentence is readily interpreted as meaning that a burglar carrying a knife threatened a student. However, (2) is in fact syntactically ambiguous between a reading where a burglar carrying a knife threatened a student, and a reading where a burglar threatened a student who was carrying a knife. To explain how we come to give preference to the first reading, we cannot rely exclusively on our standing knowledge of the grammar of English. To disambiguate phrase attachment, we must call upon commonsense reasoning and background knowledge of worldly facts (in this case, that burglars are more likely to have knives than average students), and engage in a process that many would agree is a paradigm instance of inferential work. On a similar note, consider (3).

- 3) Tim attempted to prank Mark but Susan stopped him.

In order to grasp that the correct assignment function for the free pronoun “him” should pick out Tim, rather than Mark, the listener has to exercise the due amount of sensitivity to the linguistic environment of the pronoun and carry out the relevant abductive reasoning (why would Susan stop the victim of the prank instead of its author?). Once again, this appears to be at odds with Non-Inferentialism.

Examples like (2) and (3) thus seem to provide a compelling counterargument to the thesis that listeners have a non-cognitive-inferential capacity to grasp the meaning of utterances. The task of this paper is to determine whether the Perceptual View can provide an adequate answer to the challenge raised by the context-sensitivity of language comprehension.³ The plan is as follows. Section 2 provides a complete formulation of the challenge from context-sensitivity, clarifies why it threatens the

² Earlier formulations of the challenge can be found, a.o., in Stanley (2005) and Pettit (2010). In what follows, we provide an independent formulation of the objection, and lay it out in much more detail than previous authors. For the purposes of our argument, we will use the expression “context-sensitivity” in a broad sense: in addition to paradigm cases of dependence on non-linguistic context, we will file under the general rubric of “context-sensitivity” also cases where the interpretation of an expression depends primarily on the linguistic environment where the expression occurs, and cases where the interpretation is primarily guided by commonsense reasoning and world knowledge.

³ Though of central importance, this is not the only challenge faced by the Perceptual View. For a helpful recent discussion of other challenges to the view, see Drożdżowicz (2019).

Perceptual View, and constrains the space of possible answers on behalf of the Perceptual View. Section 3 presents Brogaard's (2018; 2019) reply to the challenge: the view that in typical cases of language comprehension context-sensitivity is exercised not through inferences, but rather through top-down perceptual modulations or perceptual learning. Section 4 draws on conceptual considerations and empirical examples to argue that the exercise of context-sensitivity in language comprehension does, in fact, typically involve inference. Section 5 concludes.

2. The Objection from Context

It is widely recognized that natural language abounds with expressions and constructions whose meaning can be determined only relative to broader features of the environment, both linguistic and non-linguistic, in which they are used. Consequently, it seems safe to assume that meaning recovery is typically achieved via the exercise of similar forms of context-sensitivity as those found in (2) and (3). However, these processes seem to be paradigm cases of inferential work, and therefore threaten to invalidate the Perceptual View. Call this, for brevity, the Objection from Context. Let us lay out the objection step by step.

Objection from Context

- P1. If the Perceptual View is true, then it is not the case that the appearance of utterance meaning in language comprehension is typically the result of cognitive inferential work.
- P2. The appearance of utterance meaning in language comprehension is typically achieved by exercising various forms of context-sensitivity.
- P3. If the appearance of utterance meaning in language comprehension is typically achieved by exercising various forms of context-sensitivity, then it is typically the result of cognitive inferential work.
- P4. The appearance of utterance meaning in language comprehension is typically the result of cognitive inferential work. [P2 + P3]
- C. The Perceptual View is false. [P1 + P4]

Now that we have laid out the Objection, we can begin to examine how the Perceptual View might attempt to neutralize it. Since P4 is the combination of P2 and P3, and C is the result of the

combination of P1 and P4, there are essentially three possible ways to neutralize the Objection: by rejecting P1, by rejecting P2, or by rejecting P3.

Let us start with P1. Several aspects of this premise are worth discussing, both to avoid potential misunderstandings and to clarify the dialectic.

A first noteworthy aspect of P1 is that we interpret the Perceptual View as making a *generic* claim about how meaning recovery typically proceeds, as opposed to as an existential or universal claim. We exclude that the view should receive an existential interpretation, on which it would claim that the appearance of utterance meaning in language comprehension is not always the result of an inference. We also exclude that the view should receive a universal interpretation, on which it would claim that the appearance of utterance meaning in language comprehension is never the result of an inference. Notice that the claim that the appearance of utterance meaning is not always the result of an inference is compatible with the claim that meaning recovery is typically inferential: reinterpreted existentially, the Perceptual View would survive the Objection from Context. However, we are not interested in weaker reformulations of the view here, but rather in the view as it is currently defended.⁴

A second noteworthy aspect of P1 is that it makes clear that the Objection from Context targets only one of the conjuncts that make up the Perceptual View, viz., Non-Inferentialism (corresponding to the consequent of the conditional). One might question the relevance of this argumentative strategy. If the key philosophical issue at stake in the competition between Conservatism, Moderatism, and Liberalism is whether or not the outcome of meaning recovery is best characterized as perceptual, why focus on the conjunction of this claim with the further claim that meanings are not inferred? We have already noted that Non-Inferentialism is logically independent from Liberalism, and that it is possible to accept the liberal claim that the appearance of utterance meaning is part of perceptual phenomenology while still arguing that meaning recovery relies on inferential work. Nevertheless, there are good reasons to target specifically the conjunction of Liberalism and Non-Inferentialism. Firstly, Non-Inferentialism is philosophically interesting in its own right. Just as it is worth investigating the phenomenological correlates of meaning apprehension, it is equally interesting to investigate the nature of the process of meaning recovery, as a descriptive psychological issue that pertains to the philosophy of mind/cognitive science. Furthermore, even if Liberalism is in principle noncommittal with respect to the nature of the process of meaning recovery, examining the prospects of the Perceptual View will lead us to place constraints on the kinds of accounts of processing that can be

⁴ The generic scope is explicitly endorsed by Brogaard (2019): “[the view] is thus consistent with the *occasional* reliance on inference in order to derive the meaning of what the speaker intended to convey” (our emphasis).

embraced by Liberalism. In this sense, inquiry into the inferential vs. non-inferential nature of meaning recovery has implications for the debate between Liberalism, Moderatism, and Conservatism.

A third noteworthy aspect of P1 is that it places important restrictions on the way the advocate of the Perceptual View can answer the Objection from Context. The advocate of the Perceptual View cannot simply accept the Objection, grant that meaning recovery is typically performed inferentially, and claim that the inferential work involved penetrates perceptual appearances of meaning. While a view combining a) the claim that the outcome of meaning recovery is a perceptual appearance, b) an inferential account of meaning recovery, and c) cognitive penetration would be a consistent incarnation of Liberalism, such a view would contradict the Perceptual View's commitment to Non-Inferentialism.

A fourth, and final, noteworthy aspect of P1 is that in the expression "cognitive inferential work", the adjective "cognitive" is not redundant. As a matter of fact, one way out of the Objection would appear to be to accept the assumption that one should indeed characterize meaning recovery as typically relying on inferential work, but argue that the inferences in question occur *in perception*, rather than in cognition: meaning apprehension is inferential and perceptual, because it is performed through perceptual inferences. However, we do not think this option is attractive. To start with, while talk of "inference" is indeed common currency among, e.g., vision scientists describing the intra-modular⁵, subpersonal processes involved in early vision, this usage is considered non-literal by many philosophers⁶, including Brogaard (2019). Moreover, such perceptual "inferences", which are characteristically subpersonal, differ significantly from the type of cognitive inferences that bear on the viability of the Perceptual View. Edge detection and depth computation, for example, are done by modules in our visual system, not by *us* (the agent). Even if we granted that informationally

⁵ Unless otherwise specified, by 'module' we mean peripheral (Fodorian) modules (Fodor 1983; 2000), rather than the central modules posited by the massive modularity view (Sperber 2001). It is worth noting, in this regard, that our defense of the inferential nature of language comprehension is compatible with multiple different views concerning how specifically such an inferential capacity is implemented in our cognitive architecture (e.g., the traditional view according to which inferences are carried out by domain general central systems, as well as the massive modularity view defended by Relevance theorists). As we will argue, what matters for Inferentialism is that meaning recovery requires reason-responsive integration of wide swaths of common-sense knowledge and information about the context. This view does not contradict the possibility that the flexible way in which we typically understand discourse results from the interaction and integration of various domain-specific modules. It is neutral, in particular, on the issue of whether utterance interpretation typically involves a mind-reading module or a specialized pragmatics submodule (Wilson 2005). We thank a referee for inviting us to be more explicit on this point.

⁶ Though see Kiefer (2017), for a defense of "literal" perceptual inference.

encapsulated transitions between nonconceptual states were literally inferences, the issue at stake in the present debate is whether or not *we* infer context-dependent meanings. In this respect, the view that perceptual inferences suffice to carry out the complex information processing involved in understanding language in context is best understood as a variant of Non-Inferentialism, which faces much the same objections as those we will later raise.

Let us now turn to P2: the premise that meaning recovery typically involves context-sensitivity. One might worry that accepting P2 amounts to implicitly endorsing a form of contextualism (e.g., Recanati 2004; 2010), whereas in reality the notion that meaning recovery is systematically reliant on context is up for debate. Mightn't one reject P2 by adopting an anti-contextualist theory of linguistic meaning such as semantic minimalism (e.g., Borg 2004; 2012)? According to minimalism, competent listeners can typically recover literal meaning simply on the basis of knowledge of the lexico-syntactic properties of the language, coupled with a minimal amount of sensitivity to context.

This reply overlooks a crucial feature of the Perceptual View: that it concerns specifically the states and transitions leading to the recovery of *utterance* meaning.⁷ The Perceptual View is interested in characterizing as perceptual the conscious phenomenology of meaning apprehension, and the conscious phenomenology of meaning apprehension is about speaker-accessible utterance meaning (e.g., about a waiter's ability to spontaneously grasp, and consciously entertain, that "The ham sandwich wants a beer" means that the customer who ordered a ham sandwich also wants a beer). Importantly, the principle that speaker-accessible utterance meaning is typically affected by context is common ground between minimalists and (radical) contextualists. In frameworks like Borg's, for example, the claim that literal meaning can be recovered with a minimal amount of sensitivity to context is explicitly combined with the widely accepted view that the recovery of speaker-accessible utterance meaning is typically achieved via contextual reasoning. Because conscious meaning apprehension typically involves the latter, not the former, the idea of rejecting P2 by adopting anti-contextualism fails. Albeit implicit in the setup of the Objection from Context, this feature is worth emphasizing: it indicates that the Objection threatens the Perceptual View under any theory of meaning available on the philosophical market.

The fact that the Objection is about utterance meaning is also important for another reason, which concerns which data is, or is not, relevant to its assessment. Consider the following three empirical

⁷ By "utterance meaning", we have in mind the truth-conditional level of content expressed by sentences in their context of utterance, and corresponding roughly to Grice's notion of "what is said". Some such notion is, as far as we can tell, presupposed by all parties to the debate.

phenomena, which feature among those put forward by Brogaard (2019) in support of the Perceptual View.⁸ The first is Semantic Satiation, the phenomenon whereby repeated words gradually become meaningless for the listener. The second is the Stroop Effect, the familiar interference found when one must name the color in which a color-word is printed, as opposed to the different color which the word designates (e.g., ‘red’ printed in green characters). The third is the Pop-Out Effect, the phenomenon whereby subjects presented with a grid featuring a meaningful word (e.g., ‘telephone’) and a number of meaningless variants of that word (e.g., ‘enlehpoet’, ‘letenepho’) are able to immediately direct their attention towards the meaningful word. Each of these phenomena seems to invoke a perceptual capacity to recover the semantic properties of words, and thereby would seem to provide empirical evidence in support of the Perceptual View.

However, the matter is more complex. The first issue is that none of the empirical phenomena we have just mentioned incontrovertibly involve the perception of meaning of any sort. Listening to acoustically different repetitions of the same word (i.e., by different speakers rather than by the same speaker) does not produce semantic satiation, suggesting that the effect is actually due to adaptation to acoustic rather than semantic features (Pilotti, Antrobus, and Duff 1997). Mainstream explanations of the Stroop Effect do not all suggest that the semantic properties of the stimulus word are perceived: e.g., the interference might be due to a downstream, *post*-perceptual conflict at the response stage (for a concise survey, see MacLeod 2015). As for the Pop-out Effect, Brogaard herself concedes that the experimental data she appeals to, which is preliminary, is compatible with the possibility that the same effect would occur with any familiar string of letters, including nonsensical ones (like ‘mimsy’). Moreover, previous studies have in fact found no pop-out effect when words and non-words share low-level features (Soraci et al. 1992).

The second, more general point is that the empirical examples described by Brogaard involve *isolated* words, i.e., words presented to a subject outside the contexts in which meaning recovery typically takes place. Even if one were to grant that the meanings of isolated words are “decoded” in a very fast, automatic manner akin to perception, this would only neutralize the Objection from Context to the extent that the occasion-specific, contextually shaped meaning of utterances speakers apprehend in

⁸ We address Brogaard’s additional argument that meaning recovery is fast, automatic, and (purportedly) evidence-insensitive below. Brogaard also invokes neuro-anatomical evidence, in particular the role of Wernicke’s area in speech comprehension. We will not dwell on this particular point, except to note that the exact role of this area in speech comprehension is debated in neuroscience (e.g., Binder 2017). For critical discussion of aspects of Brogaard’s empirical arguments which fall outside the scope of our article, see Connolly (2019) and Drożdżowicz (2019).

everyday language use is derived through a similar process. As we see it, the view that the complex forms of context-sensitivity typically found in language comprehension are dealt with by processes analogous to the ones involved in decoding the meaning of isolated words requires extensive argument. Absent that, the positive evidence Brogaard provides shows at best that some inputs to the process of contextual interpretation may be perceived, but not that its outputs are. In order for empirical cases to bear on the Objection from Context, and constitute evidence in favor or against the Perceptual View, they should acknowledge P2 and feature sufficiently complex and ecologically plausible forms of context-sensitivity, rather than the mere semantic “decoding” of isolated words.

Time to recap. As we have seen, P1 simply describes the main claim of the Perceptual View, and P2 is uncontroversial. We are then left with one last option: rejecting P3, the claim that if the appearance of utterance meaning in language comprehension is typically achieved by various forms of context-sensitivity, then it is typically the result of a cognitive inference. In light of what we have said thus far, the rejection of P3 would consist in arguing that while the process of meaning recovery does typically involve various forms of processing of contextual information, most of these forms of context-sensitivity do not actually require inference. For instance, one might argue that commonsense reasoning, which is paradigmatically inferential in nature, is relatively infrequent in everyday language comprehension, and is not representative of the kind of processing involved in dealing with the most typical instances of context-dependence. If feasible, this strategy could give the Perceptual View the best of both worlds: we could acknowledge the independently plausible assumption that the exercise of context-sensitivity is pervasive and requires something more than just plain perception, without having to commit to Inferentialism about meaning recovery. It is to this way of rejecting P3 that we now turn.

3. The Top-Down Reply

The most sophisticated and powerful implementation of this strategy is found in Brogaard (2018; 2019). Her main claim is the following: the exercise of context-sensitivity in language comprehension typically results from either *top-down influences* of cognition on perception, or from *perceptual learning*. Perceptual appearances of speaker-accessible utterance meanings are typically shaped by one or the other of these non-inferential processes. In cases of top-down influence, information stored in semantic memory in a distributed fashion is implicitly retrieved and activated during online language comprehension so as to modulate meaning recovery. In cases of perceptual learning, past experience modifies the perceptual mechanisms involved in language comprehension and impacts the semantic

information hearers spontaneously pick up upon exposure to an utterance.⁹ According to such a view, top-down influences account for synchronic, short-term modulations in meaning perception, whereas perceptual learning is the source of more long-term changes. Since in both cases the processes underlying context-sensitivity shape the perceptual appearance of speaker-accessible meaning without perception receiving as input content drawn inferentially from higher-level cognition, we can reject P3, and rescue the Perceptual View from the Objection from Context. Call this, for brevity, the Top-Down Reply.

This is not the place for a comprehensive debate on the nature of inference. However, it is important to elucidate the conception of inference that underlies the Top-Down Reply. Brogaard's conception of inference, spelled out in Brogaard (2019), is clear but fairly demanding. It sets out two necessary conditions a process must satisfy in order to count as inferential.

The first necessary condition is relatively uncontroversial: inferential transitions must be *rule-governed*. Though the rules in question need not be those of valid deductive reasoning, they must be, in some broad sense, logical. The informational or representational contents of the states in an inferential transition must be related in a way that makes sense, so that each follows rationally or logically from its predecessor, even if only defeasibly so.¹⁰ This sets inferences apart from purely associative transitions. While we may associate states that follow logically from one another, there are in principle no restrictions on the range of states that can be associated via the force of habit or conditioning. So, e.g., if in one's mind thinking about lemons leads by force of habit to thinking about Mars, then we are dealing with an association, not an inference.

An empirical consequence of this condition, which Brogaard herself does not directly discuss but which has recently been stressed by Quilty-Dunn & Mandelbaum (2018), and which we take to be compatible with her view, is that the interventions that can change which inferences we are disposed to draw are different from those that affect our learned associations. Inference is reason-responsive, in the sense that providing us with new or conflicting evidence tends, all else being equal, to modify our inferential dispositions in a rational manner. Taking (2) as an example, if you previously assumed that burglars are more likely to have knives than students, but you later learn that the opposite is in fact the case, then you might cease to infer that x is a burglar and y a student from the fact that x threatened y

⁹ This is consonant with Brogaard's general view of language learning as a form of perceptual learning. See Brogaard and Gatzia (2015; 2017).

¹⁰ In saying as much, we do not mean to say, nor to attribute to Brogaard, the view that the rules must be explicitly or consciously represented by the agent performing the inference. They may, of course, be implicit in the architecture.

with a knife. By contrast, associations are not reason-responsive. One breaks an association between states not by producing a counterargument or giving a reason why one does not follow from the other, but through counterconditioning or extinction (i.e., roughly, by causing one to occur without the other). Thus, no matter how irrational one's associating, say, burglars with knives or lemons with Mars might be, being told, or proven, that burglars in fact do not typically have knives does nothing to sever the associative tie between those two concepts. This suggests an empirical test for whether or not the processes underlying the resolution of a given form of context-sensitivity are best classifiable as inferential: does providing new reasons or new evidence modify the recovered meaning in a rational manner? Or are forms of counter-conditioning or extinction required? If the former, then we are probably dealing with inference.

The second necessary condition set by Brogaard is more controversial: inferential transitions must be conscious or at least consciously accessible, i.e., such that they would be conscious in nomologically or psychologically possible counterfactual conditions. As she puts it in her (2019), this requirement on inference “restricts inferences to those that are either explicit (i.e., the subject is consciously aware of making them) or consciously accessible (i.e., the subject could – under different psychological or environmental conditions of the sort that can obtain in this world – have been aware of making them).” As we understand Brogaard, it is not meant to be merely a matter of linguistic stipulation on her part that inferences are constitutively conscious. Her goal is to cut nature at the joints, and to draw conceptual distinctions between mental states and processes that map onto real psychological kinds; she explicitly stipulates that she is interested in “‘inference’ as the phrase *ought* to be used” (emphasis ours). Hence, we take her claim that inference is best understood as necessarily conscious or consciously accessible to be open to critical evaluation.

Brogaard supports this claim as follows. Inference, she argues, is part of “personal-level” explanations of behavior. Personal-level explanations aim to make behavior rationally intelligible. To describe a mental process as inferential is to view it as evaluable in terms of whether or not it conforms to rational or logical norms. Yet, descriptions of behavior appealing exclusively to states and processes that are inaccessible to consciousness “cannot make behavior intelligible in terms of norms of rationality”. Otherwise, one could classify as intelligible in terms of norms of rationality the computations made by “unconscious machines”, which by contrast are merely causal. Hence, only conscious transitions can be characterized as genuine inferences. One consequence of this condition, Brogaard observes, is that it is impossible to characterize as inferences the transitions that occur within

sub-personal (Fodorian) modules, such as the computational transitions that take place in early visual processing.

We are happy to grant the (common, though not always clear) distinction between personal and subpersonal levels.¹¹ We are also happy to grant that the transitions occurring within Fodorian modules, such as the computational transitions of early visual processing, should not (for present purposes) be characterized as inferences.¹² However, we are skeptical of Brogaard’s assumption that ruling out subpersonal (e.g., intra-modular or “subdoxastic” in the sense of Stich 1978) processing as not genuinely inferential should lead us to conclude that no unconscious processing is characterizable as inferential. Brogaard notes that if a state or process is subpersonal, then it is not consciously accessible. But the converse implication, for which Brogaard does not provide an argument, fails: there are personal-level states and processes that are unconscious. In fact, there are reasons – both empirical and conceptual – to hold that there are genuinely personal-level unconscious inferences, as Quilty-Dunn & Mandelbaum (2018) argue at some length. Among the many cases they describe, one example in point is cognitive dissonance, which involves non-associative, rational, personal-level connections between states (e.g., instantiated patterns of *modus tollens*), which are nonetheless not consciously accessible to the subject. Following this and other work in social psychology and the psychology of reasoning (e.g., Johnson-Laird 2008), it seems therefore more plausible to us to distinguish several subspecies of inference, some conscious and others not, rather than require that *all* inferences be conscious – which has the unwelcome consequence of failing to capture deep psychological commonalities between processes that differ only along the conscious/unconscious dimension.¹³

¹¹ For useful clarifications of the personal/subpersonal distinction, see Drayson (2014) and Lyons (2016). Lyons (2016: 250-251) lists several reasons why “subpersonal assumptions” are to be contrasted with those attributable to the agent, including that the agent needn’t have the concepts involved to specify their content (e.g., zero-crossing), and that they do not tend to conflict with other beliefs of the agent, nor with other intra-modular assumptions.

¹² We believe that the primary reason why intra-modular transitions should not be characterized as inferences is that they are characteristically informationally encapsulated. Hence, that they are not reason-responsive, which we have noted is a plausible necessary feature of inference. E.g., giving a subject reasons to judge that the arrows in the Müller-Lyer illusion are of equal length does not make the illusion disappear. Note that transitions within *central* modules, if there are such, may count as personal-level by our criteria. Indeed, an advantage of the moderately permissive view of inference we defend is that it would be strange if the massive modularity view led to the conclusion that, since much of cognition is modular, nearly no inferences ever occur. By contrast, it seems this could be an unwelcome consequence of adopting a more restrictive view of inference like Brogaard’s.

¹³ The same inference might be conscious in a strong sense (i.e., such that its premise- and conclusion-states are conscious and the subject endorses the transition between them, forming the judgment that the conclusion follows from the

One might nonetheless have the following worry.¹⁴ As we noted, Brogaard adopts the restrictive notion of inference we have described because she wishes inferences to play a certain epistemological role, viz., feature in personal-level explanations of behavior. By adopting a more liberal notion, are we not ignoring this aspect of the debate, and do we not risk talking at cross-purposes? First of all, like Brogaard in her (2019), we are interested in evaluating the psychological claims of the Perceptual View “on epistemically neutral grounds”: the issue at stake in the debate is the descriptive psychology of meaning recovery, not its epistemology. As a consequence, it seems dialectically fair to background epistemological issues, and refrain from enforcing any specific interpretation of the notion of inference based on epistemological desiderata. Moreover, on a more substantive note, it is far from obvious that the epistemological considerations staged by Brogaard support restricting the notion of inference to conscious processes. As recently argued by Lyons (2016), it does not follow from the mere fact that a belief is “psychologically immediate”, i.e., simply occurs to us, that it is therefore necessarily *epistemically* noninferential, i.e., that it is immediately, basically justified (if justified at all). It seems plausible to say, instead, that when a belief is based on personal-level unconscious reasoning from unconscious beliefs, these unconscious beliefs serve as evidence for the psychologically immediate one, and bear on its justificatory status (Lyons 2016: 247). In other words, drawing inspiration from Lyons, we can grant Brogaard’s requirement that the notion of “inference” should inform personal-level explanations of behavior, as well as her point that subpersonal inferences do not satisfy that requirement, but at the same time maintain that *unconscious* inferences, when (and only when) they are personal-level, also satisfy the requirement. At any rate, it is the opposite assumption that requires argument.

We thus obtain the following less stringent characterization of the nature of inference (which is not meant as a definition): inference is a personal-level transition between mental states which a) is not necessarily conscious, and b) distinguishes itself from other transitions between mental states (e.g., associations) in virtue of its rule-governedness and reason-responsiveness. As we have made clear, though this conception of inference is not as restrictive as Brogaard’s, it does not render our disagreement with her merely verbal, nor is it obviously unsuited to epistemological concerns. With

premises), or conscious in a weaker sense (i.e., such that its premise- and conclusion-states are conscious but the transition between them is not consciously endorsed), or completely unconscious. For example, see Reverberi et al. (2012) for empirical evidence that *modus ponens* inferences can be carried out unconsciously.

¹⁴ We are grateful to an anonymous referee for raising this worry.

this more encompassing characterization in sight, we can now examine whether the Top-Down Reply offers an effective response to the Objection from Context.

4. The Revenge of Context

As a benchmark against which to test whether the Top-Down Reply provides an effective response to the Objection from Context, we will assume that the reply is successful if and only if two specific conditions are met. These two conditions, which simply spell out the main commitments of the Top-Down Reply, are as follows.

Availability Condition

Information about context is available to the perceptual processing of utterances.

No Inference Condition

The exercise of context-sensitivity in meaning recovery does not typically require inferential work.

The Availability Condition states the simple requirement that information about context be available to perceptual processing. Insofar as the Perceptual View claims appearances of speaker-accessible meanings to be typically perceptual, and we accept that such appearances typically require the exercise of context-sensitivity, it must allow that the relevant informational resources are available to perception. The No Inference Condition corresponds to the requirement that the processes by which contextual information affects the appearance of utterance meaning cannot typically involve cognitive inferential work (given the characterization of inference we spelled out in Section 3).

Let us examine each of these two conditions in turn, starting with the Availability Condition. We believe this condition is, in fact, supported by some empirical evidence. Different lines of research converge on the claim that perceptual (especially auditory) processing in language comprehension is, in various ways, shaped by information about context. Skipper (2014) shows that the activation of the auditory cortex appears to be inversely correlated with the degree of meaningfulness or relevance of speech sounds: the more linguistically salient the sounds are, the less active the auditory cortex is, and vice versa. Özyürek (2014) provides evidence for a tight integration between vocal and visual channels in processing, and shows that the level of integrated processing can be modulated by the communicative context (e.g., co-speech gestures are processed together with speech and give rise to a single multi-modal percept employing non-auditory information sources). The point extends to the

integration of facial information (Riedel et al. 2015). Kelly et al. (2007) demonstrate that our brain integrates speech and gesture less strongly when the two modalities are perceived as not intentionally coupled (i.e., gesture and speech are produced by two different persons) than when they are perceived as produced by the same person. This result also demonstrates that information about the intentional relationship between gesture and speech modulates neural processes during the early integration of the two modalities. Finally, Hartwigsen et al. (2015) show that in the auditory processing of acoustically degraded but highly predictable sentence endings, predictive information about the upcoming word aids auditory processing completing the perceptual representation of the speech signal. See, e.g., (4).

- 4) Lois had to read lips because she was [*degraded pronunciation of 'deaf'*].

This is by no means an exhaustive survey of the empirical literature in this area. But we can grant that it suffices to tip the balance in favor of the Availability Condition: some information about context can in fact be accessed by the perceptual processing of utterances.

Now let us move to the No Inference Condition. It states that the Top-Down Reply provides a successful answer to the Objection from Context if and only if, in meaning recovery, context-sensitivity is typically exercised without inferential work. To determine whether or not this condition is satisfied, we shall focus on three additional examples of context-sensitivity, in which meaning recovery seems to require precisely the kind of inferential work barred by the No Inference Condition.

The first case involves polysemy. In many instances of sentence interpretation, competent listeners disambiguate the meaning of words simply on the basis of the semantic properties of their neighboring constituents. Consider (5) and (6).

- 5) The book was originally written in Chinese.
6) The book weighs 15.5 ounces.

Most current treatments of polysemy, both theoretical (see Ortega-Andrés and Vicente 2019) and experimental (e.g., Frisson 2015), assume that nouns like “book” denote a complex semantic object ambiguous between an abstract entity, as in (5), and a physical item, as in (6). The selection of the specific interpretation to be assigned to the noun is derived by selecting the sense that matches the type of the property designated by the attached verb phrase (“was originally written in Chinese” vs. “weighs

15.5 ounces”). Even when broader context does not aid the interpretation in any way (e.g., the sentence is uttered in isolation rather than as part of a larger discourse, and there are no other elements priming either of the two word senses), listeners are consistently able to recover the appropriate interpretation of the noun. Crucially for our purposes, however, these operations exhibit the familiar symptoms of *bona fide* cognitive inference: there is a coherent, logically intelligible chain of steps between the inspection of the semantic type of the two possible senses of the noun, the evaluation of the type requested by the verb phrase, and the selection of the sense that satisfies type-mismatch constraints.

The second example concerns the recovery of pronoun-antecedent relations in intersentential anaphora. Consider the simple sentence pair in (7).

- 7) John invited Mark over for dinner. He spent hours cleaning the house.

In this example, the free pronoun “he” has two possible antecedents, “John” and “Mark”, but is readily interpreted as referring to John. Unlike (5) and (6), listeners cannot reach this conclusion simply by inspecting the linguistic context where the pronoun occurs. Instead, they have to compute the statistical likelihood of the two competing antecedents relative to a prior background of world knowledge. This is the received wisdom encapsulated by most current formalized models of free anaphora resolution. For example, Sagi and Rips (2014) claim that on null-context exposure to (7), speakers conclude that the antecedent of “he” is “John” by engaging in implicit causal reasoning about the identity of the person who is more likely to have cleaned the house. Among other things, they: a) form an implicit hypothesis about what sort of property should be instantiated by the referent of the pronoun (being most likely to have cleaned the house); b) access background world knowledge that in social situations like dinners the person playing the host-role, rather than the person playing the guest-role, will instantiate that property; c) analyze the linguistic material in the first sentence, which indicates that John issued the invitation and that the dinner was held at John’s place, to conclude that John played the host-role; d) determine that, as a result of his playing the host-role, John instantiates the property relevant to the interpretation of the pronoun. Notice how easily we can invert the outcome if we replace the second sentence with a clause providing evidence for the rival interpretation.

- 8) John invited Mark over for dinner. He spent hours picking the perfect gift.

Similar comments apply to (9)-(10) and (11)-(12) (inspired by Wilson 1992).

- 9) The police arrested the demonstrators because they feared violence.
- 10) The police arrested the demonstrators because they committed violence.

- 11) Tom was beating someone for no reason. The man must be out of his mind.
- 12) Tom was beating someone for no reason. The man must be badly hurt now.

The third example concerns cases where an anaphoric pronoun without any candidate antecedent is provided an implicit antecedent by the preceding discourse. Consider (13) (from Cornish 1999).

- 13) A: “Why didn’t you write to me?”
B: “I did. Well, I tried to. But I always tore them up.”

The relevant element in this case is the unstressed pronoun “them” in B’s response, which is readily interpreted as referring to a series of letters even if letters are never explicitly mentioned in the dialogue. While in (7) the task of the interpreter consisted in picking between two candidate overt antecedents, (13) adds a further layer of complexity: there is no overt antecedent. The referent of the pronoun, rather, appears to be derived from the illocutionary point of A’s initial question, which bears on the expectation that A and B should have engaged in written correspondence and on the non-existence of the correspondence in question, and from the lexical semantics of “tore [...] up”, which indicates that the vehicle of the correspondence was physical and hence presumably consisted of letters (rather than, say, emails). (13) is yet another example of the extent to which inferences based on existing discourse representations, linguistic knowledge and commonsense reasoning are mobilized in the interpretation of context-sensitive expressions. Similarly, see (14), where the reference of “they” appears to be jointly established by an implicit inference that the pronoun is intended to pick out the authors of “Sympathy for the Devil”, and by encyclopedic knowledge that the song “Sympathy for the Devil” was written by the Stones.

- 14) Do you remember Sympathy for the Devil? Well, they made another album in 2016.

See also (15), where the referent of “the thing” seems to require an even more complex plausibility inference: the object at stake has to fit in a pocket and can be used to check a Twitter account, hence it is in all likelihood a smartphone.

15) I tried to check my twitter, but the thing just fell out of my pocket and broke.

As we have seen, examples (5)-(15) raise important concerns about the idea that the No Inference Condition might provide explanatory purchase on the processing labor typically taking place during the online treatment of context-sensitivity. Consistent with our reconstruction of the Top-Down Reply in Section 3, we see three main ways the advocate of the Perceptual View might respond to the challenge.

The first reply would be to grant that examples like (5)-(15) are incompatible with the No Inference Condition but claim that they are overall infrequent, and certainly not typical. Because of their rarity, examples like (5)-(15) may well be inconsistent with the No Inference Condition, but they fail to disprove the generic claims of the Perceptual View. We can make two simple considerations against this reply. First, there is no obvious reason to place the burden of proof on the typicality of the cases themselves, instead than on the claim that the cases are exceptions to the rule of the No Inference Condition. As it stands, the Perceptual View provides no explicit reasons to think that these cases are atypical, and arguing that they are simply because this would rescue the No Inference Condition would be question-begging. Second, we believe that, taken together, (5)-(15) *do* form a corpus of cases which is sufficiently heterogeneous and robust to constitute an explanandum for a generic account of meaning recovery like the Perceptual View. Moreover, the corpus can be expanded easily, in the same spirit as the examples we provided. For example, notice how a simple substitution of the verb of the second sentence leads speakers to interpret (16) as a form of narration (the events represented by the second sentence succeed those represented by the first) and (17) as a form of explanation (the events represented by the second sentence have caused those represented by the first).

16) Max fell. Susan helped him.

17) Max fell. Susan pushed him.

Crucially, the variation in this case affects the very core of “what is said” by the sentence pairs at hand, and each preference seem to require implicit reasoning and inference-making about the most likely

discourse structure of the pair. Cases involving implicit processing of discourse structure are pervasive in language comprehension (see, a.o., Kehler 2002; Asher and Lascarides 2003; Kehler et al. 2008; Ginzburg 2012). Arguably, they provide a more realistic setting for testing theories of meaning recovery than textbook examples of isolated utterances or utterance pairs. So it stands to reason to include the processing of discourse effects like (16) and (17) among the explananda of the Perceptual View.

The second reply would be to grant that examples like (5)-(15) are typical but maintain that they can in fact be accounted for through *perceptual learning*. In essence, this would amount to claiming that we should consider characterizing the bodies of knowledge and the processes underlying the forms of context-sensitivity found in (5)-(15) as processing constraints wired and taking place within the perceptual system. We think that this rejoinder faces a major difficulty. While it seems reasonable to characterize as perceptual learning the acquisition of some aspects of linguistic competence, such as the ability to spontaneously perceive utterances as organized into syllables, words, phonetic segments and other interesting discrete units, it is much less clear, and indeed very unlikely, that morphological knowledge, semantic knowledge, and especially the various contextual skills underlying our ability to interpret (5)-(15) can be similarly characterized.¹⁵ For example, the proponent of the Perceptual View would have to produce a model of the acquisition, storage and recruitment of syntactic knowledge which does without the received view that they consist of bodies of symbolic information irreducible to simple constraints on perceptual processing. They would have to characterize as learned constraints on perceptual processing the body of implicit information about binding conditions (see Büring 2005) that triggers judgments of ungrammaticality on exposure to sentences featuring a pronoun bound within its clause, such as (18).

18) * Rebecca_i hit her_j on the head with a zucchini.

Unless such an account is formulated in sufficient detail, the idea that the Perceptual View offers a viable account of the way knowledge of the language is acquired and stored in the mind/brain of competent speakers and contributes to meaning apprehension is, we fear, more of a promissory note than a hard-earned empirical theory.

The third reply would be to grant that examples like (5)-(15) are widespread in language, grant that the examples cannot be accounted for through perceptual learning, and insist that despite initial

¹⁵ Connolly (2019: 154-178) develops this argument at length.

appearances they nevertheless do not involve inferential work since they involve top-down influences. To start, notice that it is not sufficient to argue, as Brogaard comes close to doing, that no genuine inference is involved because the relevant processing is at least partly unconscious. We grant that, in typical instances of language comprehension, what we are conscious of is simply the output of the process of meaning recovery (that “the book” in (5) refers to something abstract, that “he” in (8) refers to Mark, that “they” in (9) refers to the police), rather than the process of meaning recovery itself. However, in keeping with the characterization of inference we have laid out in Section 4, this does not show that the prior processing leading to this interpretation is non-inferential. One might, of course, attempt to characterize the disambiguation of “the book” in (5) vs. (6) as occurring through associations between each of the relevant disambiguations and the accompanying predicates. Our rejoinder is that logical polysemy resolution involves logical or rational connections – as do the various other examples we have provided. For instance, the connection in (13) between figuring out that the speaker is referring to the letters and tearing is most plausibly abductive: this interpretation makes good sense (as opposed to other interpretations that may involve equally strong associations with tearing – say, muscles in the mind of an athlete).

At this point, mightn’t the defender of the Perceptual View point out that just because there is a rational connection between states, this does not conclusively show that they are not *also* associated? After all, premises and conclusions of arguments one frequently runs through presumably come to be associated in memory. And since any state can be associated with any other in principle, it is difficult to conclusively prove that any given transition is inferential, even when it involves rational connections between states. For instance, one might argue that what we are doing is really providing rational reconstructions of processes that as a matter of empirical fact are carried out associatively.¹⁶

Our response is twofold. First, the more a task involves drawing a novel or creative connection, the less attractive a purely associative account becomes. This is precisely a feature of many of our examples, which would be interpreted by a competent speaker in the ways we have discussed even if she were encountering them for the first time. We do not deny that, in practice, linguistic performance may often involve routines, heuristics, defaults – in sum, various associative shortcuts. Nonetheless, we are skeptical that this observation can be used to neutralize our plea for inferences. A truly linguistically competent speaker would be capable of producing the interpretations we have described even on first exposure to these sorts of examples. Insofar as we accept that a theory of language comprehension should predict our capacity to recover the meaning of *novel* sentences in our language,

¹⁶ We thank a referee for pushing this point.

which need not conform to prior expectations, and insofar as we recognize that such cases cannot easily be covered by previous associative links, the idea that the examples we have provided might be explained associatively loses traction.¹⁷

Second, to the extent that one grants the psychological reality of the inference/association distinction, as do all parties to the debate, there *is* a way, if not to conclusively establish that a process is inferential, at least to strongly suggest that it is. Although, e.g., reference resolution for expressions dependent on linguistic and/or non-linguistic context is typically fast, automatic, and often unconscious, its *reason-responsive* nature strongly suggests that it is nonetheless inferential. In examples (7)-(12), a relevant change in the linguistic or extra-linguistic environment leads to a corresponding change in recovered meaning. This responsiveness to evidence is precisely what one would expect if pronoun interpretation were the result of a typical inference. Counter-conditioning or extinction seems to have no effect on the appearance of meaning. But changes in evidence of, e.g., the speaker's referential intention do. Appearances of utterance meaning are therefore quite unlike the prototypical cases of evidence-insensitive perceptual appearances found, e.g., in the Müller-Lyer illusion.

The two points we have just made are worth stressing in connection with the account of the resolution of pronominal reference provided by Recanati (2004: 30-31) which could, at first glance, seem to contradict our inferential account. Recanati suggests that examples like (7)-(12) can be accounted for in an accessibility-based framework, involving processes that are not properly inferential, by resorting to the concept of a "frame" (in the sense of Fillmore 1982). For example, hearing (8) might automatically evoke in the hearer a "dinner invitation" frame, which represents the various roles involved in a typical situation of that sort. In (8), Mark is linked to the role of invitee (rather than host) in the first sentence, making him a more accessible candidate for being the referent of "he" in the second sentence.

¹⁷ In the field of Artificial Intelligence, for example, pairs similar to (7)-(12) above are known as 'Winograd Schemas'. They are currently used as an alternative to the Turing Test to test machine comprehension (Levesque, Davis and Morgenstern 2012). Tellingly, the 'Winograd Schema Challenge' is trivial for humans, but very hard for computer programs (Davis, Morgenstern, Ortiz 2017). Levesque et al. note that this is because succeeding above chance requires one to "figure out what is going on", as they nicely put it. Winograd Schemas are "Google-proof": access to a huge corpus of text (which actual humans typically lack) and powerful statistical methods don't help much. To the extent that these methods bear some resemblance to processes of associative learning, the difficulty of designing a program that comes at all close to human performance on the Winograd Schema Challenge using such techniques provides a practical illustration of the challenges facing any associative account.

In reply, we again note, in keeping with our point about novelty, that this sort of account fails to explain cases where fitting individuals into roles within stereotyped scripts or schemas is not what enables us to resolve reference. Carston (2007) argues as much by providing an example in the spirit of the following. Suppose I meet a colleague at the university who tells me: “Bob is back from sabbatical”. As it happens, I know two Bobs, both university professors: Bob₁, who I am very close to and have recently encountered, and who is highly accessible, and Bob₂, who is just a distant acquaintance, but who I know to be known to my colleague. Despite the fact that nothing in the frame (say, the university setting) supports a Bob₂ rather than Bob₁ interpretation, and despite the fact that Bob₁ is otherwise more accessible, I correctly understand my interlocutor to mean Bob₂, contrary to the accessibility account. I simply draw on my knowledge that my interlocutor knows Bob₂, and not Bob₁.

Moreover, we maintain, in keeping with our point about reason-responsiveness, that even cases involving frames or schemas are often best characterized as inferential. Frames are a way to organize information within our conceptual belief system, and which help us efficiently reason by avoiding combinatorial explosion (Thagard 1984). Frames may allow contextual factors to feature in interpretation without requiring hearers to deploy dedicated mind-reading abilities and mental state meta-representation¹⁸. Nevertheless, the processes that, according to Recanati, involve frames remain characteristically reason sensitive, and involve reasonable transitions between fully conceptual states. In our view, it makes better sense to talk in such cases of frame-based inference, i.e., to see frames as cognitive devices for efficiently *implementing* reasoning, as opposed to viewing processing involving frames as non-inferential. Viewing frames as associative seems to us to be a byproduct of the assumption, which we have argued is mistaken, that if a process is fast and unconscious (as the processes of accessing a frame and using it to resolve reference may well be) then it is subpersonal and associative. Once we accept the existence of genuinely unconscious inferences, there is little obstacle to including frame-based reasoning among these.

Which leads us to the final moral. We have suggested that the Top-Down Reply would offer a viable response to the Objection from Context on behalf of the Liberal View if and only if two specific conditions were met: the Availability and the No Inference Condition. We have argued that while the Availability Condition appears to be supported by some empirical evidence, a more careful consideration of the processing labor involved in central instances of context-sensitivity gives strong reasons to think that the No Inference Condition is not satisfied. We therefore conclude that there is a

¹⁸ For more on the distinction between pragmatic processes that involve mental state attribution and those that do not, see Kissine (2016).

strong case to be made that in favor of the view that the exercise of context-sensitivity in meaning recovery does typically involve inferential work. In some cases, this inferential work is performed consciously or can be consciously accessed through a suitable introspective effort. In many cases, this inferential work takes place outside the spotlight of our conscious awareness. However, as we have emphasized, this is no reason to refrain from classifying it as genuinely inferential.

The point is symptomatic, in our view, of a broader but fundamental difference between perception and linguistic understanding. Despite the popularity of phrases like “the testimony of the senses”, determining what is, say, visually present in one’s surroundings is usually not a matter of trying to cooperate with another person, whereas linguistic communication is. Only the latter activity is by essence one in which we deploy personal-level rational capacities and responsiveness to evidence provided by another rational being to navigate our social environment, coordinate with our peers, justify our beliefs and actions to others (see Mercier & Sperber 2019). Viewed from this standpoint, the notion that recovering the meaning of utterances typically involves inferential work should come as no surprise.

5. Conclusion

This paper has argued that the Top-Down Reply to the Objection from Context fails. The challenge raised by context-sensitivity to the Perceptual View, understood as the combination of Liberalism about the phenomenology of meaning apprehension with Non-Inferentialism about the process of meaning recovery, therefore remains unanswered. We have provided empirical evidence and conceptual considerations in favor of the view that the exercise of context-sensitivity in meaning recovery is, in fact, typically achieved thanks to unconscious, personal-level inferential work.

Where does this leave us with respect to the debate between Liberalism, Moderatism and Conservatism, which we introduced at the beginning of the paper? Because the Perceptual View is only one possible incarnation of Liberalism, our argument does not provide any direct objection against Liberalism *toto coelo*: the Semantic Contrast we introduced at the beginning of our discussion may still turn out to be a difference in perceptual phenomenology regardless of the truth of our thesis that we should accept Inferentialism about meaning recovery. However, our case for Inferentialism does place an important argumentative burden on Liberalism: proponents of Liberalism must now pursue their claim that appearances of utterance meaning are typically perceptual within the broader view that meaning recovery is typically cognitive-inferential in nature.

Such a position is not in principle untenable. However, the conjunction of Liberalism and Inferentialism seem to imply the view that perceptual appearances of utterance meaning are typically shaped by inferences, and hence typically involve cognitive penetration. The nature and the existence of cognitive penetration are thorny subjects which we have studiously avoided in this paper. On minimal assumptions, however, committing to cognitive penetration would generate two significant costs for Liberalism. First, the plausibility of Liberalism would hinge on the independent debate about whether cognitive penetration *ever* occurs (in language comprehension or elsewhere; see, e.g., Firestone and Scholl 2016). Second, if appearances of utterance meanings were penetrated by the workings of cognitive-inferential mechanisms, Liberalism might lose one of the attractive features that, according to its proponents, should lead us to consider it in the first place: its purported ability to avoid the familiar epistemological worry that we cannot really “know” the meaning of the utterances we hear. If our auditory experiences of meanings were cognitively penetrated, they would display the very epistemological characteristics (such as defeasibility) that gave rise to such a worry in the first place.

References

- Asher, N. and A. Lascarides. 2003. *Logics of Conversation*. Cambridge: Cambridge University Press.
- Balcerak Jackson, B. 2019. Against the perceptual model of utterance comprehension. *Philosophical Studies* 176: 387-405.
- Bayne, T. 2009. Perception and the reach of phenomenal content. *The Philosophical Quarterly* 59: 385-404.
- Binder, J. R. 2017. Current controversies on Wernicke's area and its role in language. *Current Neurology and Neuroscience Reports* 17: 58.
- Borg, E. 2004. *Minimal Semantics*. Oxford: Oxford University Press.
- Borg, E. 2012. *Pursuing Meaning*. Oxford: Oxford University Press.
- Brogaard, B. 2017. The publicity of meaning and the perceptual approach to speech comprehension. *ProtoSociology* 34: 144-162.
- Brogaard, B. 2018. In defense of hearing meanings. *Synthese* 195: 2967-2983.
- Brogaard, B. 2019. Seeing and hearing meanings: A non-inferential approach to speech comprehension. In *Inference and Consciousness*, ed. T. Chan and A. Nes. London: Routledge.
- Brogaard, B. and D. E. Gatzia. 2015. Is the auditory system cognitively penetrable? *Frontiers in Psychology* 6: 1166. doi: 10.3389/fpsyg.2015.01166
- Brogaard, B. and D. E. Gatzia. 2017. The real epistemic significance of perceptual learning. *Inquiry* 61: 543-558.
- Büring, D. 2005. *Binding Theory*. Cambridge: Cambridge University Press.
- Carston, R. 2007. How many pragmatic systems are there? In María José Frápolli (ed.), *Saying, Meaning and Referring: Essays on François Recanati's Philosophy of Language*. London: Palgrave-Macmillan.
- Connolly, K. 2019. *Perceptual Learning: The Flexibility of the Senses*. New York: Oxford University Press.
- Cornish, F. 1999. *Anaphora, Discourse and Understanding*. Oxford: Oxford University Press.
- Davidson, D. 1973. Radical interpretation. *Dialectica* 27: 313-328.
- Davis, E., Morgenstern, L., and Ortiz, C. L. 2017. The first Winograd schema challenge at IJCAI-16. *AI Magazine*, 38(3), 97-98.
- Drayson, Z. 2014. The personal/subpersonal distinction. *Philosophy Compass* 9: 338-346.

- Drożdżowicz, A. forthcoming. Do we hear meanings? – between perception and cognition. *Inquiry*, DOI: 10.1080/0020174X.2019.1612774.
- Fillmore, C. 1982. Frame Semantics, in The Linguistic Society of Korea (ed.), *Linguistics in the Morning Calm*, Hanshin: 111-138.
- Firestone, C., and B. J. Scholl. 2016. Cognition does not affect perception: Evaluating the evidence for ‘top-down’ effects. *Behavioral and Brain Sciences*, e229: 1-77.
- Fodor, J.A. 1983. *The Modularity of Mind*. MIT Press.
- Fodor, J.A. 2000. *The Mind Doesn't Work That Way: The Scope and Limits of Computational Psychology*. MIT Press.
- Frisson, S. 2015. About bound and scary books: The processing of book polysemies. *Lingua* 157: 17-35.
- Ginzburg, J. 2012. *The Interactive Stance: Meaning for Conversation*. Oxford: Oxford University Press.
- Hartwigsen, G., T. Golombek, and J. Obleser. 2015. Repetitive transcranial magnetic stimulation over left angular gyrus modulates the predictability gain in degraded speech comprehension. *Cortex* 68: 100-110.
- Isac, D., and C. Reiss. 2013. *I-Language: An Introduction to Linguistics as Cognitive Science*, 2nd edn. Oxford: Oxford University Press.
- Johnson-Laird, P. 2008. *How We Reason*. Oxford: Oxford University Press.
- Kehler, A. 2002. *Coherence, Reference, and the Theory of Grammar*. CSLI Publications.
- Kehler, A., Kertz, L., Rohde, H., & Elman, J. 2008. Coherence and coreference revisited. *Journal of Semantics* 25: 1-44.
- Kelly, S. D., S. Ward, P. Creigh, and J. Bartolott. 2007. An intentional stance modulates the integration of gesture and speech during comprehension. *Brain and Language* 101: 222-233.
- Kiefer, A. 2017. Literal Perceptual Inference. In T. Metzinger & W. Wiese (Eds.). *Philosophy and Predictive Processing*: 17. Frankfurt am Main: MIND Group. Doi: 10.15502/9783958573185.
- Kissine, M. 2016. Pragmatics as metacognitive control. *Frontiers in Psychology* 6: 20-57.
- Levesque, H., Davis, E., & Morgenstern, L. (2012, May). The winograd schema challenge. In *Thirteenth International Conference on the Principles of Knowledge Representation and Reasoning*.
- Longworth, G. 2008. Linguistic understanding and knowledge. *Nous* 42: 50-79.

- Ludlow, P. 2011. *The Philosophy of Generative Linguistics*. Oxford: Oxford University Press.
- Lyons, J. C. 2016. Unconscious evidence. *Philosophical Issues* 26: 243-262.
- MacLeod, C. 2015. The Stroop Effect. *Encyclopedia of Color Science and Technology*. DOI 10.1007/978-3-642-27851-8_67-1.
- Mercier, H., and D. Sperber 2019. *The Enigma of Reason*. Cambridge, MA: Harvard University Press.
- O’Callaghan, C. 2011. Against hearing meanings. *The Philosophical Quarterly* 61: 783-807.
- Ortega-Andrés, M., and A. Vicente. 2019. Polysemy and co-predication. *Glossa: A Journal of General Linguistics* 4. doi: <https://doi.org/10.5334/gjgl.564>.
- Özyürek, A. 2014. Hearing and seeing meaning in speech and gesture: insights from brain and behaviour. *Philosophical Transactions of the Royal Society B* 369: 20130296.
- Pilotti, M., Antrobus, J. S., & Duff, M. (1997). The effect of presemantic acoustic adaptation on semantic “satiation”. *Memory & Cognition* 25(3): 305-312.
- Pettit, D. 2010. On the epistemology and psychology of speech comprehension. *The Baltic International Yearbook of Cognition, Logic and Communication* 5: 1-43.
- Prinz, J. 2006. Beyond appearances: The content of sensation and perception. In *Perceptual Experience*, eds. T. S. Gendler and J. Hawthorne. 434-459. Oxford: Oxford University Press.
- Quilty-Dunn, J., and E. Mandelbaum. 2018. Inferential transitions. *Australasian Journal of Philosophy* 96: 532-547.
- Recanati, F. 2004. *Literal Meaning*. Cambridge: Cambridge University Press.
- Recanati, F. 2010. *Truth-Conditional Pragmatics*. Oxford: Clarendon Press.
- Reiland, I. 2015. On experiencing meanings. *The Southern Journal of Philosophy* 53, 4: 481-492.
- Reverberi C., Pischedda D., Burigo M., and Cherubini P. 2012. Deduction without awareness. *Acta Psychologica* 139: 244-253.
- Riedel, P., P. Ragert, S. Schelinski, S. Kiebel, and K. Kriegstein. 2015. Visual face-movement sensitive cortex is relevant for auditory-only speech recognition. *Cortex* 68: 86-99.
- Sagi, E., and L. Rips. 2014. Identity, causality, and pronoun ambiguity. *Topics in Cognitive Science* 6: 663-680.
- Siegel, S. 2006. Which properties are represented in perception? In *Perceptual Experience*, ed. T. Szabo Gendler and J. Hawthorne. 481-503. Oxford: Oxford University Press.
- Skipper, J. 2014. Echoes of the spoken past: how auditory cortex hears context during speech perception. *Philosophical Transactions of the Royal Society B* 369: 20130297.

- Soraci, S. A., Franks, J. J., Carlin, M. T., Hoehn, T. P., & Hardy, J. K. 1992. A “popout” effect with words and nonwords. *Bulletin of the Psychonomic Society* 30(4): 290-292.
- Sperber, D. 2001. In defense of massive modularity. In E. Dupoux (ed.), *Language, Brain and Cognitive Development: Essays in Honor of Jacques Mehler*. Cambridge, MA: MIT Press.
- Stanley, J. 2005. Hornsby on the phenomenology of speech. *The Aristotelian Society Supplementary Volume 79*: 131-146.
- Stich, S.P. 1978. Belief and subdoxastic states. *Philosophy of Science* 48: 499-518.
- Strawson, G. 1994. *Mental Reality*. Cambridge, MA: MIT Press.
- Thagard, P. 1984. Frames, knowledge, and inference. *Synthese* 61: 233-259.
- Tye, M. 2000. *Consciousness, Color, and Content*. MIT Press, Cambridge, MA.
- Wilson, D. 1992. Relevance and reference. *UCL Working Papers in Linguistics* 4: 167-191.
- Wilson, D. 2005. New directions for research on pragmatics and modularity. *Lingua* 115: 1129-1146.