Synthese (2018) 195:2967–2983
https://doi.org/10.1007/s11229-016-1178-x

In defense of hearing meanings

Berit Brogaard[1]

Received: 26 January 2016 / Accepted: 26 July 2016 / Published online: 6 August 2016

Abstract According to the inferential view of language comprehension, we hear a speaker's utterance and infer what was said, drawing on our competence in the syntax and semantics of the language together with background information. On the alternative perceptual view, fluent speakers have a non-inferential capacity to perceive the content of speech. On this view, when we hear a speaker's utterance, the experience confers some degree of justification on our beliefs about what was said in the absence of defeaters. So, in the absence of defeaters, we can come to know what was said merely on the basis of hearing the utterance. Several arguments have been offered against a pure perceptual view of language comprehension, among others, arguments pointing to its alleged difficulties accounting for homophones and the context-sensitivity of ordinary language. After responding to challenges to the perceptual view of language comprehension, I provide a new argument in favor of the perceptual view by looking closer at the dependence of the justificatory qualities of experience on the notion of a defeater as well as the perceptual nature of language learning and language processing.

Keywords Ambiguity · Cognitive penetration · Cognitive phenomenology · Language comprehension · Perceptual learning · Phenomenal contrast argument · Phenomenal dogmatism · Presentational phenomenology · Polysemy · Top-down influences

✉ Berit Brogaard
brogaardb@gmail.com

[1] Department of Philosophy, University of Miami, Coral Gables, FL, USA

1 Introduction

According to the inferential view of language comprehension, we hear a speaker's utterance and infer (likely unconsciously but not necessarily on a subpersonal level) what was said, drawing on our competence in the syntax and semantics of the language together with background information. On the alternative perceptual view, fluent speakers of a language have a non-inferential capacity to auditorily perceive not just the sounds of speech but also its content.[1] On this latter view, when we hear a speaker's utterance, the experience confers some degree of justification on our beliefs about what was said in the absence of defeaters (Fricker 2003).[2] So, in the absence of defeaters, we can come to know what was said merely on the basis of hearing the utterance. A defeater is best understood as a personal-level belief that either undercuts or rebuts an existing belief (Pollock 1986, 1987; Politzer and Bonnefon 2006). Rebutters act by directly weakening the likelihood that the belief in question is true. For example, you see that people are wet when entering the department and come to believe that it's raining but a reliable witness informs you that it's not raining. Your belief may still be true but it is now less likely to be true. Undercutters, by contrast, attack the connection between the evidence one has for a belief and the belief. They weaken the sufficiency of the evidence for the core belief to be true. For example, you form a perceptually-based belief that a table is red but then discover that the table is illuminated by red light. The table may still be red, so your belief may be true but it is now less likely to be true. In both cases, the prima facie justification is undermined and your belief therefore is not justified.

One problem with the inferential view of language comprehension is that it lends itself to the view that we cannot come to know the meanings of utterances, or the even more radical view that meanings themselves are indeterminate.[3] As Pettit (2010) has argued, the problem arises owing to the epistemic implications of the inferential view for language acquisition. Acquiring a correct belief set about the syntax and semantics of the language doesn't suffice for comprehending what speakers say; those beliefs would also need to be justified. This is the idea underlying Davidson's (1973) thought experiment of the radical interpreter who is faced with the task of interpreting a completely foreign language. The theorist interprets the language on the basis of the available evidence, which are observations of the linguistic and non-linguistic behavior of the speakers. The problem is that the available evidence underdetermines facts about meaning, since that evidence is compatible with rival hypotheses about the meaning

[1] It is slightly misleading to talk about *auditorily perceiving what is said* (or loosely: *hearing meanings*). In far the most cases, we don't auditorily perceive what is said, but see people say something. The latter case is not a case of auditory experience as such but rather one of multisensory experience. It is to be expected, of course, that seeing lip movement and gestures can contribute in significant ways to our perceptual grasp of what is said. I shall set aside these more complicated cases here but hope to deal with them in future work.

[2] The second epistemic component is, in principle, an optional addition to the perceptual view. However, the attractiveness of the perceptual view may in part depend on the cogency of the argument for the epistemic component.

[3] Here I follow the tradition in linguistics of using 'utterance contents' and 'utterance meanings' synonymously.

of the utterances. This leads Davidson to argue not merely that we cannot come to know the meaning of the utterances we hear but also that there are no determinate facts about meaning, because facts about meaning cannot outstrip the ability of all the speakers of a language.

This sort of skepticism does not follow from the perceptual view of language comprehension, because on the latter view, our experiences of what is said immediately justify our beliefs about what is said without any reliance on further belief or theorizing, at least in the absence of defeaters. So, in the absence of defeaters, our beliefs about the meanings of utterances are justified.

The perceptual view thus has an advantage over the inferential view insofar as it can block a potential route to a radical view of language comprehension (cf. McDowell 1978, 1981).[4] If the perceptual view is at all plausible, it is for this reason strongly preferable to the inferential view. The problem is that the perceptual view may not be plausible.

Two main arguments have been offered against the hypothesis that the perceptual view explains the phenomenal contrast between listening to familiar and foreign languages. One turns on the alleged difficulties accounting for cases of homophones, words that sound the same but have different meanings, such as 'pole' and 'poll' (O'Callaghan 2011). The other turns on an alleged inconsistency of the perceptual view with the pervasive context-sensitivity of ordinary language (Stanley 2005; Pettit 2010).

Here I will defend the explanatory role of the perceptual view against these potentially devastating objections. After responding to the arguments against the explanatory role of the perceptual view, I provide a positive argument in its favor.

My defense of the perceptual view is grounded in two empirically validated hypotheses about language and perception. (i) Language learning is largely a type of perceptual learning, and (ii) language comprehension is a kind of perception that is massively influenced by top-down factors. The influence of top-down factors on language comprehension, I will argue, does not undermine the role of appearances of what was said as immediate justifiers of belief. Whereas top-down influences by themselves do not have this negative effect, top-down influences capable of cognitively penetrating appearances do. But, as we will see, when appearances are cognitively penetrated, they do not have the qualities needed to serve as immediate justifiers. The justificatory qualities of appearances reside in their resistance to defeaters—a property that cognitively penetrated appearances do not have. Defeaters thus play a crucial role in the formulation of phenomenal dogmatism. They help us distinguish between the appearances that provide prima facie justification for belief from those that are unable to do so. Considerations of speech comprehension will help illustrate this point. Or so I will argue.

[4] There are, of course, other ways to block the Davidsonian line of argument. So, this line of argument should not be taken to be the main reason to adopt the perceptual view of language comprehension.

2 The argument from homophones

The auditory perceptual experience of listening to speech in known and in unknown languages are remarkably different. In both cases we arguably hear the same sounds uttered but our experiences differ in phenomenology (Bayne 2009; Siegel 2005; Pettit 2010; O'Callaghan 2011; Reiland 2015a). One difference between listening to speech in a known and an unknown language is that we understand the known language. So, a natural suggestion is that the difference in phenomenology derives from the meaning of the utterance.

O'Callaghan (2011) offers an objection to this argument. While he agrees that there is a difference in phenomenology between listening to speech in known and unknown languages, he argues against the claim that the phenomenal contrast is best explained in terms of our auditory perception of meanings. The phenomenal contrast is better understood in terms of ability to differentiate language-specific phonological properties of the known language. Although this hypothesis is consistent with the view that we can non-inferentially perceive the content of speech, this explanation of the phenomenal contrast between listening to known and unknown languages indicates that O'Callaghan would argue against the view that we have an auditory sensory phenomenology of meanings.

Against the hypothesis that the phenomenal contrast between listening to a known and an unknown language is best explained by the hypothesis that we can hear meanings, O'Callaghan brings up the case of homophones. Homophones are expressions that are pronounced the same way but differ in meaning. They include homonyms, which share a spelling, as in the case of 'bank' (financial institution) and 'bank' (river bank), and heterographs, which do not share a spelling but are nonetheless pronounced the same way, such as 'pole' and 'poll'. Two utterances of the homophones 'pole' and 'poll' involve acoustically identical sounds. When listening to those sounds, we do not detect any difference in phenomenal character, despite the potential difference in meaning.

This is the gist of O'Callaghan's case against the view that the phenomenal contrast between listening to known and unknown languages is best explained in terms of an auditory experience of meanings. When the argument is understood in this way, however, it does not establish the hypothesis that the phenomenal contrast is not best explained perceptually (cf. Reiland 2015a). Procedural/functional words, such as 'but', 'or', 'and' and 'not' can perform a multiplicity of functions in the language. For instance, the sentence 'Otavio is in his office, and he is writing' is true just in case both conjuncts are true. 'Otavio might be in his office, and he might be in Brazil', on the other hand, is true just in case one of the conjuncts is true. The general linguistic meaning of procedural/functional words, such as 'but', 'or', 'and' and 'not' is normally inaccessible to consciousness. But so is the lexical meaning of descriptive words, such as 'chair', 'healthy' and 'door' because of the massive polysemy of ordinary language (Recanati 2004). 'Chair', for instance, could mean (among many other things) a seat for one person, the head of a department or organization, and an office of position or authority. Even if one could come up with some exhaustive disjunctive lexical entry specifying the multiple linguistic functions of procedural and descriptive words, it is implausible to think that we ordinarily comprehend speech by consciously accessing

such complex lexical entries. If, however, linguistic meaning elutes consciousness, it can be neither perceived nor grasped.

What is accessible to consciousness is the utterance (or occasion) meaning of a particular use of a word. We have no trouble accessing the meaning of an utterance of the sentence 'the patient is healthy but she is still not eating anything healthy', despite the different occurrences of the highly polysemous word 'healthy'. The utterance meaning we comprehend in individual instances are what we use as evidence when theorizing about linguistic meanings. Linguistic meanings are thus in some sense theoretical constructs.

These considerations point to a problem with O'Callaghan's argument, if construed as an argument against experiencing semantic properties.[5] The sounds associated with the words 'pole' and 'poll' are indeed identical and the words mean different things. But simply pronouncing these words isn't a case of *using the words*. It is merely a case of pronouncing the sounds associated with two different lexical entries. Yet, as already argued, we cannot hear the meaning of sounds associated with lexical entries because the meanings of lexical entries are ineffable; they are not ordinarily accessible to consciousness. So, the simple argument from homophony doesn't work.

O'Callaghan preempts something like this objection but proceeds by arguing that the lack of difference in phenomenology persists even when we focus on particular uses of the words 'poll'/'pole'. He invites us to listen to utterances of 'Ernest used the pole to vault over the high bar', 'Last year Mac visited the southern pole of Earth', and 'Bubb won the greatest number of votes in our latest poll'. O'Callaghan maintains that even when uttered as part of a sentence, we will be unable to attend to anything audible in the three utterances of 'poll/pole' that makes them different.

It seems, however, that the specific meanings of the homophones *do* make a difference to the phenomenology of the listening experiences. If the same sounds ('poll'/'pole') appeared in a foreign language, as in the case of an utterance of the Danish sentence 'Giv dukken til Poll',[6] the experience of the word would be different. We would have no impression of experiencing a meaning. In fact, O'Callaghan's own explanation of the phenomenal difference between listening to a known language and a foreign language, which we will revisit below, has exactly the same alleged problematic implication.

Having rejected that an auditory experience of meanings can explain the phenomenal contrast between listening to a known language and listening to a foreign language, O'Callaghan owes us a different explanation of the phenomenal contrast. His explanation turns on familiarity with the sounds of a language. Learning a language, he argues, changes the temporal and qualitative features which speech sounds are experienced as having. When we learn a language we become better at detecting the language-specific phonological properties of the language, which alters how the sounds are experienced.

[5] This is not to say that O'Callaghan's intention in putting forth the argument was to establish that we don't perceive semantic properties but only that one might potentially use this sort of argument to attempt to show that we don't perceive semantic properties.

[6] Here we can imagine that someone is simply named 'Poll', pronounced like the English word. The utterer would then be asking someone to give the doll to Poll.

There are several problems with this explanation. First, O'Callaghan's explanation of differences in phenomenology between familiar and foreign languages cannot explain the feeling that there is an immediately perceptible difference between different "in-context" utterances containing different homophones.

Second, as we will see below, two different languages can have exactly the same speech sounds but nonetheless have different meanings associated with those speech sounds. Even though the speech sounds are the same, learning the language nonetheless changes the phenomenology of our overall experiences of utterances in the language. The most plausible explanation of this difference is a shift in the perception of what was said.

Third, the debate about whether one can hear the meaning of utterances carries over to written language. There is, arguably, a phenomenal difference between looking at a message written in a known language and a message written in a foreign language. If we are immediately aware of the content of the message, then this could plausibly be taken to explain the difference in phenomenology. But if we merely see the configurations of the letters of the message and then go through a step of inferences to reach an interpretation, then a different explanation of the difference in phenomenology between seeing a message written in a known language versus a foreign language is called for. Changes in the temporal and qualitative features that speech sounds seem to have cannot explain the phenomenal difference in this case, and, as we will see below, it is questionable that learning a language can change the qualitative features which graphemes are experienced as having without us also having an auditory experience of what the message communicates.

3 The argument from context sensitivity

Another argument that has been set forth against the view that we can be directly perceptually aware of what is said and that the resulting perceptual representations can confer immediate justification on our beliefs about what was said turns on the pervasiveness of context sensitivity in ordinary language. Jason Stanley formulates the objection as follows:

> Those who hold that language understanding is akin to some kind of non-inferential perceptual grasping face the obvious objection that the pervasive context sensitivity and ambiguity of natural language sentences forces hearers to engage in inferential reasoning about meaning in order to grasp what is said by an utterance. When someone utters the sentence 'The policeman arrested the robber. He was wearing a mask', we generally interpret the pronoun 'he' as referring to the robber, rather than the policeman. We arrive at this interpretation by exploiting inferences about the plausibility of interpreting the pronoun in different ways, inferences guided by our knowledge of meaning together with background knowledge about the world. Virtually every sentence we hear contains context-dependent expressions. Therefore, virtually all of our experience as language interpreters involves making consciously accessible linguistically guided inferences about semantic content. (2005, pp. 131–132)

Stanley's point is that we need to make inferences by exploiting our background knowledge in order to arrive at an interpretation of pronouns, such as 'he', 'she' and 'it.' But since context-sensitivity is so pervasive in ordinary language, it would seem that we cannot have knowledge of what the speaker said through direct perceptual awareness of meanings. This suggests that the perceptual view is mistaken.

Pettit (2010), who is favorable toward the perceptual view, argues that the pervasive context sensitivity of speech suggests that speech comprehension is not modular in Fodor's (1983) sense. A type of processing is modular just if it is not subject to cognitive influences but is 'informationally encapsulated'. Since the pervasive context sensitivity of speech indicates that the processing of speech is affected by cognitive influences, this massive context sensitivity indicates that speech comprehension is not modular.

Pettit proceeds to argue that speech comprehension, and linguistic competence more generally, is not 'warrant apt', it is not the sort of thing that can be immediately justified. The reason for this, he argues, is that the processing of language that leads to language comprehension is sub-personal. Unlike belief, such processes or states are inaccessible to consciousness, which do not make them apt for being warranted.

Pettit's point about the failure of speech comprehension to be modular does indeed provide an ample line of defense against Stanley's objection. We may well be directly perceptually aware of speakers' intended meanings because background information influences the appearances of meaning through top-down processes. These top-down influences need not be instances of cognitive penetration in Pylyshyn's (1999) sense. For a higher-level cognitive state or process to cognitively penetrate a lower-level state or process, there must be a semantically-coherent chain of steps that begins with the cognitive state and eventually results in an alteration of the lower-level state (Brogaard and Chomanski 2015). For example, if I think that there is a pink elephant in the room and this thought affects my visual system in such a way that it visually seems to me that there is a pink elephant in the room, then my visual state is cognitively penetrated. If, on the other hand, my difficulties comprehending a talk on migraines result in anxiety that in turn gives me migraine auras, then my vision is not cognitively penetrated but is altered through a top-down influence, viz., anxiety. The distinction between top-down influences and cognitive penetration is crucial to the epistemic dimension of the perceptual view. I shall return to this in a subsequent section.

The suggestion that speech comprehension depends on background information as a result of top-down processes is compatible with it sometimes being the result of direct perceptual awareness of meanings (Brogaard and Gatzia 2015). Stanley's point about the pervasive context-sensitivity of speech thus does not present a challenge to the perceptual view. The main challenge to the perceptual view, as we will see, turns on the justificatory role of appearances of speaker meanings.

4 The justificatory qualities of experience and the notion of a defeater

There has been a lot of debate about whether visual seemings just are visual experiences (Ghijsen 2015; Chudnoff and DiDomenico 2015) or whether they are distinct from such experiences (Tucker 2010; Lyons 2015; Conee 2013; Brogaard 2013a; Bergmann

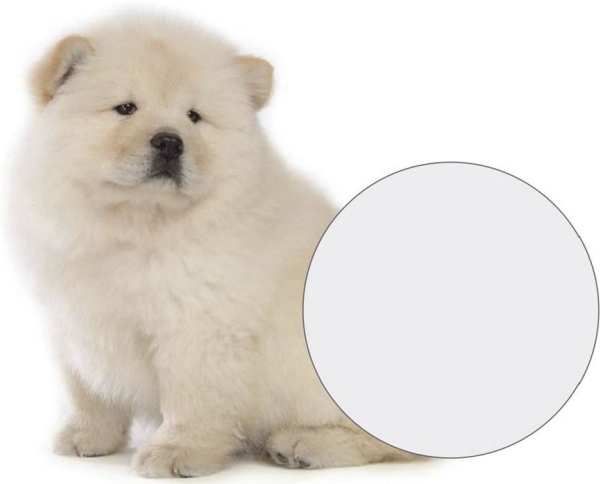

Fig. 1 The *dog* is partially occluded. The truth-conditions for your appearance of the *dog* include both the proposition that that is a *dog* and awareness of the truth-maker for that proposition but it does not include awareness of the *dog*'*s tail*. So, while your experience of the *dog* has presentational phenomenology, your experience that the occluded parts are parts of a *dog* does not

2013; Reiland 2015b). A number of thinkers have argued that experiences by their very nature are low-level, such as visual experiences of color, extension and texture or auditory experiences of pitch and timbre (e.g., Reiland 2015b), but the question here concerns high-level phenomena that are on a par with appearances of emotions or personality (see Brogaard in press). As I am not taking a stance on this issue in this paper and perceptual awareness of meanings evidently is a higher-order phenomenon, I will henceforth use the term 'appearance' (or 'seeming') rather than 'experience'.

According to the phenomenal dogmatist, at least some perceptual appearances provide immediate justification for belief. But not all appearances do, at least not all by themselves. Consider the illustration in Fig. 1.

If it appears to me that what is hidden behind the occlusion is part of a dog, and I come to believe it, this appearance does not by itself justify my belief. At best, it confers justification on my belief together with background assumptions about dogs and tails.

What quality do those appearances have that are in a position to confer immediate justification of belief?

Elijah Chudnoff has argued that only appearances with a presentational phenomenology confer immediate justification on belief, that is, phenomenal dogmatism should be restricted to those cases in which appearances, or seemings, have presentational character, such as the appearance of the whole dog 'popping' out in front of our eyes. As Chudnoff puts it:

> One principled way to restrict phenomenal conservatism, then, is to restrict it to those propositions with respect to which seemings have presentational character: whenever it seems to you that p and your seeming has presentational character

> with respect to p, then you thereby have at least prima facie justification for believing that p. If it seems to you that p and your seeming lacks presentational character with respect to p, you still might have prima facie justification for believing that p, but, as the cognitive penetration cases suggest, it will depend in part on background information (Chudnoff 2014; cf. Chudnoff 2013, pp. 90, 94; Chudnoff 2016; Chudnoff forthcoming).

To a first approximation, seemings have presentational character only when their accuracy conditions 'include both *p* and awareness of a truthmaker for *p*' (Chudnoff 2016).

Returning to the occluded dog, your visual experience of the dog makes you aware of the proposition that the dog is sitting as well as the truthmaker for that proposition, but it does not make you aware of a truthmaker for the proposition that the dog has a short tail, a long tail or no tail or that the tail continues in one direction rather than another. On this view, the content of an experience of the sitting dog is not simply *the dog is sitting* but something like: *that dog is sitting*, *and it seems that I am aware of a truthmaker for the proposition that that dog is sitting*. Experiences of occluded parts of objects have a different content. For example, an experience of the occluded part of the dog being part of a dog might have the content: *that part is part of that dog but I am not aware of a truthmaker for the proposition that that part is part of that dog*.

Chudnoff's proposal, however, runs into trouble with respect to experiences of what is said by sound sequences. The trouble is that auditory sequences give rise to the illusion of auditorily experienced meanings that appear to be evidence insensitive in just the same way as lower-level visual illusions (cf. Longworth 2008).[7] YouTube booms with videos of cats and dogs who allegedly can say adorable things such as 'I love you'.[8] The experience that the cat or dog said *I love you* is remarkably resistant to any defeaters, and there are plenty. We all know that most of these video recordings are recordings of sounds that happen to resemble the sounds of an utterance of 'I love you'. Even if cats and dogs can be trained to utter the sounds 'I love you', the sounds do not have the utterance meaning we seem to hear. Cats and dogs do not have the linguistic competence of an English speaker.

Another example of an illusion of hearing contents comes from the artificial language Food Tongue, created by math campers in Northern America in 2004.[9] The constituents of Food Tongue are food words in English, such as "cherry pie," "hot dog" and "kiwi." But the food words don't retain their ordinary English meaning, and the language contains unique grammatical rules for how you can combine them. Food Tongue usually follows normal English word order (subject–verb–object, preposition–noun, etc.). Everything else (such as tense, prefix, number, and person) is done with

[7] Longworth (2008) notes that the appearance that a sentence like 'More people have been to France than I have' is meaningful may persist even after we realize that it is, in fact, incomprehensible.

[8] See, e.g., https://www.youtube.com/watch?v=adAJ3y4EL9I for a cat allegedly expressing how much she loves her owner. Retrieved on 15 December 2015.

[9] Food Tongue Wiki, http://foodtongue.soy/.

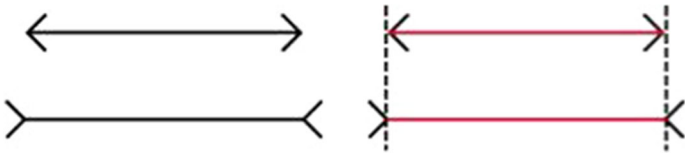

Fig. 2 The Müller–Lyer Illusion. Even when you learn that the *line segments* on the *left* have the same length, they continue to appear as if they have different lengths

separate words, such as 'yellow-pepper', 'spinach', 'eggs', and 'pear'.[10] Consider the following Food Tongue sentences:[11]

(1)
(a) Yogurt! Plantain wasabi! Apple cauliflower Berry.
(b) Apple oyster Food Tongue; apple sauce Food Tongue grass ham-sandwich tongue.
(c) Food Tongue grass tongue quiche stew camper fish ham-sandwich yellow-pepper dough. Mint tongue-slice calamari grass food.

If you were to hear random utterances of these sentences without having learned any Food Tongue, you would get the impression that someone was attempting to commit their grocery list to short-term memory, not that they were actually saying something about languages and math campers. The illusion of simply hearing random listings of English words for foods persists even when we possess the knowledge that the speaker has intentionally expressed meaningful propositions.

These illusions of hearing meanings are evidence resistant in exactly the same way as lower-level visual illusions, such as the Müller–Lyer Illusion (Fig. 2). It perceptually seem to me that the two lines in the Müller–Lyer Illusion have different lengths even when I know that they do not.

As Pylyshyn (1999) points out, this sort of insulation against rational influences is a mark of perceptual experience (or appearance) proper. The fact that auditory appearances of what was said also possess this mark suggests that this type of auditory appearance may be the kind of appearance that can confer immediate justification on belief. Yet appearances of what was said by an utterance do not possess a presentational phenomenology, in Chudnoff's sense. This is because, on his view, appearances with a presentational phenomenology make us aware of a truth-maker of the content of the appearance. Yet the truth-maker for 'what was said is *p*' includes entities that are not immediately, or genuinely, perceptible, for instance knowledge of grammar and compositionality, linguistic conventions and background information about the speaker. So, an auditory appearance that a speaker said that *p* does not make us imme-

[10] "FoodTongue: Interview with Alan Huang", http://lemmingsblog.blogspot.com/2009/05/foodtongue-interview-with-alan-huang.html. Retrieved on 16 December 2015.

[11] 1(a) Yogurt! Plantain wasabi! Apple cauliflower Berry [Hello. And welcome. My name is Berit], 1(b) Apple oyster Food Tongue; apple sauce Food Tongue grass ham-sandwich tongue [I love Food Tongue; I think Food Tongue is a math camper language], 1(c) Food Tongue grass tongue quiche stew camper fish ham-sandwich yellow-pepper dough. Mint tongue-slice calamari grass food. [Food Tongue is a language that math campers invented in mathcamp in the past. Each lexical item is a kind of food.]

diately aware of the truth-maker for 'what was said is *p*'. Hence, auditory appearances of what was said do not have a presentational phenomenology.

An alternative proposal is that the appearances that serve as immediate justifiers of beliefs are those that are evidence resistant. Call these appearances 'phenomenal'. We can distinguish between phenomenal and epistemic appearances as follows. Phenomenal appearances are those that persist, even in the presence of defeaters, at least in the case of perceivers with a typical psychology. Phenomenal appearances confer immediate justification upon belief in virtue of their resistance to evidence.

5 Perceptual learning and top-down influences

As noted above, Stanley (2005) believes that cases of context-sensitivity in the language makes it the case that appearances of what was said can justify belief only in conjunction with background information. I have already raised some issues with this argument, inspired by Pettit. A further concern is that the learning that is required in order to skillfully detect phonemes in a new language and match them to meanings is a kind of *perceptual learning*.

Contextual background information is not something we combine with auditory appearances in order to comprehend what is said. Background information directly influences the formation of visual and auditory appearances. One way it does that is in the process of perceptual learning. In perceptual learning our sensory system alters in a way that affects how things appear to us.

Consider the case of expert chess players. Whereas novices are only able to encode the position of the individual chess pieces in long-term memory, expert chess players encode chess configurations. The basic unit encoded in long-term memory is the 'chunk', which consists of a configuration of pieces that are frequently encountered together and that are related by type, color, role, and position (Chase and Simon 1973a, b). The number of figurations that the expert player has stored in long-term memory can be as high as 300,000 (Gobet and Simon 2000). The chunks can also be encoded in a combined form known as 'templates' (Gobet and Simon 1996).

Language learning evidently also proceeds via perceptual learning. During the initial stages of second language learning, for example, speakers use controlled processes with focal attention to task demands (McLaughlin et al. 1983). They may consciously employ grammatical rules when producing sentences and use translation when reading. At more advanced stages automatic processes are employed, and the attention demands decrease.

When learning to read a new language, the brain transitions from a process of recognizing words as random strings of letters to a process of visually representing them in chunks, where a 'chunk' can be considered a kind of visual object. This was shown in a study where researchers recruited a group of college students to learn the meaning of 150 nonsense words (Glezer et al. 2015). Before they learned the meaning of the words, their brain registered them as a jumble of symbols. But after they learned their meaning, their brain dedicated a circuit of neurons to each word in the visual word form area in the temporal cortex, an area which stores a visual representation of known words. The results indicate that once the volunteers learned to comprehend

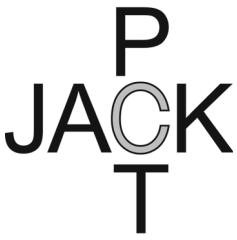

Fig. 3 The *middle letter* appears as a *C* when it occurs in 'Jack' and as an *O* when it occurs in 'pot', owing to the holistic processing of words

the words, they began to see them as units rather than sets of random letters, making it possible to read and comprehend at a faster pace.

A note of clarification is here in order. Despite the fact that particular objects, letters and words come in indefinitely many physical varieties, they can be recognized as belonging to a certain type of object, letter or word. For instance, the word 'cat' can be typed in different fonts or hand-written in a multiplicity of styles and yet be recognized as the word 'cat'. In Fig. 3, for instance, readers immediately recognize the two words as 'POT' and 'JACK', despite the shared middle grapheme.

Owing to this sort of variation in the physical appearance of words and letters, the brain doesn't dedicate a neural circuit to each instance of an object, letter or word. Rather, the brain responds to a limited number of distinctive features (Gordon 2004). For example, the distinctive features for the letter R include a vertical line, a closed curve and a diagonal line, whereas those for the letter O include only a closed curve. When the brain transitions to a holistic way of processing through the process of perceptual learning, these circuits coding for these features come to form a single circuit that fire together when the stimulus is present.

The second way background information directly influences the formation of visual and auditory appearances is via top-down processing. Background information concerning what the speaker is likely to convey guides the processing of the visual or auditory input and the resulting visual or auditory appearances of what was said. Consider the following ambiguous newspaper headings (Matlin 2013, p. 306):

(2)
(a) Eye drops off the shelf.
(b) Squad helps dog bite victims.
(c) British left waffles on Falkland Islands.
(d) Bombing Rocks Hope for Peace.
(e) Clinton wins budget; more lies ahead.
(f) Miners refuse to work after death.
(g) Kids make nutritious snacks.
(h) Local high school dropouts cut in half.
(i) Iraqi head seeks arms.
(j) Oklahoma is among places where tongues are disappearing.

Although we do occasionally need to stop and think about what the intended meaning is, in many cases our expectations at a higher level of processing automatically influences lower-level processing, quickly generating an appearance of the intended

meaning. 2(a), for instance, is often immediately comprehended as conveying that a certain type of eye drops has been taken of the shelves, owing to safety issues.

Background information guides the appearance of what the speaker's intended meaning was in other ways as well. It is a psychological fact that we don't actually fully process information we read or hear. Consider:

(3) After the plane crash at the border of Chile and Argentina, the authorities needed to decide where to bury the survivors.

On a quick read, (3) does not strike us as if anything is wrong with it. It does not strike us as startling or weird. More careful attention reveals that it would be exceedingly odd for the authorities to bury people who *survived* a plane crash.

Average college students can read about 255 words per minute, which would be an impossible feat if they were to stop and think about the meaning of every single word. Fernanda Ferreira and her colleagues have proposed that we comprehend language by using a partial, or "good-enough," approach (Christianson et al. 2010; Ferreira et al. 2002; Swets et al. 2008). On this approach we process only part of what we read or hear and fill in the rest through top-down processing. In the case of (3) we do not fully process the word 'survivor' and fill in the meaning of 'passenger' instead, perhaps partially on the grounds that few people ever survive plane crashes.

In this respect comprehending language is not unlike visual experience. Most visual experiences of objects are experiences of occluded objects. We don't see the backside of objects or the parts covered by trees, tables and car doors. Likewise, we don't hear all the meaningful information presented by utterances. In both cases, however, the objects and the meanings communicated are presented as units in experience with top-down processing aiding in filling in the missing information.

The pervasiveness of top-down influences on perceptual appearances of meaning gives rise to a potential challenge for phenomenal dogmatism. If visual or auditory seemings are cognitively penetrated, then it would seem that there are certain bootstrapping cases that threaten to undermine phenomenal dogmatism. Consider the following case from Siegel (2012). Mary believes unjustifiably that John is mad at her. When she sees John, who is in fact behaving very friendly toward her, her unjustified belief cognitively penetrates her visual system and modifies her perception of John. Accordingly, it comes to visually seem to her that he is mad at her. According to unrestricted phenomenal dogmatism, however, this visual seeming confers immediate justification upon her belief that he is mad at her. Her unjustified belief that he is mad at her thus ends up justifying itself.

Siegel's case rests on the plausibility that beliefs can cognitively penetrate the visual system. Let it be granted for argument's sake that this assumption is correct. In previous work (Brogaard 2013b) I have argued that the case doesn't present a problem for phenomenal dogmatism. The reason for this is that only genuine phenomenal seemings can serve as immediate justifiers, whereas epistemic seemings cannot. If the seeming in question is phenomenal rather than epistemic, then it would persist in the presence of a defeater, if Mary has a typical psychology. If someone were to point out to Mary that John shows no signs of anger, then it would no longer seem to her that John is mad at her. Defeaters thus can help us distinguish between the appearances that provide prima facie justification for belief from those that are unable to do so and avoid bootstrapping cases.

It is straightforward to conjure up analogous bootstrapping cases for the case of speech comprehension. Suppose I have the background information that you always complain about the quality of departmental talks at the wine and cheese reception following it. After Dr. Brown's talk we chat at the reception. This one time you express your admiration of the talk but my background information modifies my perception of what you said. So, I hear it as a complaint about the quality of the talk. In this case, my belief that you were going to say something negative about the talk makes it seem to me that you said something unfavorable about the talk.

This case, like Siegel's case, does not threaten to undermine phenomenal dogmatism, however. If the seeming can serve as an immediate justifier of my belief, it would not persist if I were presented with a defeater. For instance, if you recorded your exact utterance and played it back to me (and my psychology is typical), it would no longer seem to me that you were saying something unfavorable about the talk. But if the appearance has those properties, then on my view it cannot serve as an immediate justifier of belief. Hence, the bootstrapping case does not threaten to undermine phenomenal dogmatism about what was said.

More generally speaking, perceptual appearances that are cognitively penetrated are not among the mental states that can serve as immediate justifiers of belief. Genuine phenomenal appearances of speaker meanings can serve as immediate justifiers of belief, even though they depend on background information through top-down processes. Since those top-down processes are not normally cases of cognitive penetration, they present no challenge to the view that appearances of utterance meanings can confer immediate justification on beliefs.

6 Conclusion

I have argued that we can become directly aware of speaker meanings through auditory perception, and that our knowledge of what was said owes primarily to the role of perceptual appearances as immediate justifiers of belief. This is the perceptual view of speech comprehension. This view stands in opposition to the inferential view. On the latter view, we rely on inferential processes in order to determine what speakers say.

Two potentially serious objections have been presented to one of the main considerations in favor of the perceptual view, viz., that there is phenomenal difference between listening to familiar and foreign languages. One is that in the case of homophones, such as 'pole' and 'poll', our auditory experiences are identical in spite of the fact that homophones differ in meaning. My reply to this objection turned on the distinction between linguistic meaning and what is said by the utterance. Linguistic meaning evades consciousness and therefore is not something that can ordinarily be consciously perceived and not something that speakers ordinarily think about. Only utterance (or occasion) meanings are ordinarily accessible to consciousness but if we consider genuine cases in which homophones are used in sentences uttered in context, then the auditory phenomenology associated with the perception of what is said are different for the different uses.

It should be emphasized that the perceptual appearances that can give us direct access to speaker meanings are appearances brought about by speech fragments, not

pronunciations of individual words. Evidence for this comes from cases in which we fill in unheard or ignored gaps in speech through top-down influences. Unheard or unattended words do not have a phenomenology. So, even if the phenomenology of the appearances of individual homophones (e.g., 'poll' and 'pole') in distinct utterances were the same, this would not undermine the plausibility of contrast argument for the perceptual view. What matters is the phenomenology of the whole, which inevitably will be different in cases where there is any speech comprehension at all.

The other objection to the thought that the perceptual view best explains the phenomenal contrast between listening to familiar and foreign languages turned on the pervasive context sensitivity of ordinary language. It may be thought that if speech comprehension depends on background information in order for the speaker to grasp the speaker's intended reference of a context-sensitive word, then comprehension can result only after inferences have been made from what was heard together with the background information the listener possesses. This, however, is not the case. Research suggests that language comprehension is subject to top-down influences that aid in generating a perceptual appearance of the message conveyed by the speaker. So, context-sensitivity is perfectly compatible with the idea that we can be directly perceptually aware of speaker meanings.

My positive case for the perceptual view was grounded in two empirically supported hypotheses about language comprehension. One was that language learning is a kind of perceptual learning that changes how we perceive utterances.

The other was that language comprehension proceeds via processing in the auditory or visual system aided by top-down influences. Background information is not the only type of top-down influence that the perception of utterance meanings is subject to. Studies suggest that the functional architecture of perceptual processing involves primarily top-down modulation (Brogaard and Gatzia 2015). Top-down influences exerted throughout the auditory systems include (among other things): memory, attention, (prior) knowledge of syntax or words and experience-based expectations pertaining to the speaker's accent, gender, and vocal folds or tract (Suga et al. 2002; Chandrasekaran et al. 2013; Gilbert and Li 2013; Lotto and Holt 2011).

As I have argued elsewhere, these top-down influences are not cases of cognitive penetration (Brogaard and Gatzia 2015). Appearances of speaker meanings could, in principle, be subject to cognitive penetration. However, as I have argued here, appearances of this kind are not the kinds of mental state that can serve as immediate justifiers of belief, because they are too easily defeated. Appearances that can serve this role are those that persist even in the presence of defeaters, under assumptions of minimal hearer rationality. Hence, on the view proposed here phenomenal dogmatism—the view that perceptual appearances can confer immediate justification upon belief in the absence of defeaters—can be preserved, even for the case of speech comprehension.

Acknowledgements I am grateful to Brendan Balcerak-Jackson, Ned Block, Anna Drożdżowicz, Casey O'Callaghan, Francois Recanati, Josh Weisberg and Wayne Wu for helpful discussion of these issues and to Elijah Chudnoff, Kathrin Glüer, Anandi Hattiangadi, Casey Landers, Luca Moretti, Peter Pagin, Tommaso Piazza, David Poston, Dag Westerståhl, audiences at Stockholm and Houston and two anonymous reviewers for this journal for helpful comments on a previous version of the paper.

References

Bayne, T. (2009). Perception and the reach of phenomenal content. *The Philosophical Quarterly*, *59*(236), 385–404.

Bergmann, M. (2013). Phenomenal conservatism and the dilemma for internalism. In C. Tucker (Ed.), *Seemings and justification: New essays on dogmatism and phenomenal conservatism* (pp. 154–180). Oxford: Oxford University Press.

Brogaard, B. (2013a). Do we perceive natural kind properties? *Philosophical Studies*, *162*(1), 35–42.

Brogaard, B. (2013b). Phenomenal seemings and sensible dogmatism. In C. Tucker (Ed.), *Seemings and justification: New essays on dogmatism and phenomenal conservatism* (pp. 270–289). New York: Oxford University Press.

Brogaard, B. (in press). Perceptual appearances of personality. *Philosophical Topics*.

Brogaard, B., & Chomanski, B. (2015). Cognitive penetrability and high-level properties in perception: Unrelated phenomena? *Pacific Philosophical Quarterly, 96*(2015), 469–486.

Brogaard, B., & Gatzia, D. E. (2015). Is the auditory system cognitively penetrable? *Frontiers in Psychology*, *6*, 1166. doi:10.3389/fpsyg.2015.01166.

Chandrasekaran, B., Skoe, E., & Kraus, N. (2013). An integrative model of subcortical auditory plasticity. *Brain Topography*. doi:10.1007/s10548-013-0323-9.

Chase, W. G., & Simon, H. A. (1973a). The mind's eye in chess. In W. G. Chase (Ed.), *Visual information processing* (pp. 215–281). New York: Academic.

Chase, W. G., & Simon, H. A. (1973b). Perception in chess. *Cognitive Psychology*, *4*, 55–81.

Christianson, K., Luke, S. G., & Ferreira, F. (2010). Effects on plausibility on structural priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*, 538–544.

Chudnoff, E. (2013). *Intuition*. Oxford: Oxford University Press.

Chudnoff, E. (2014). Review of Tucker (eds.) *Seemings and justification*. *Notre Dame Philosophical Reviews*. http://ndpr.nd.edu/news/48118-seemings-and-justificationnew-essays-on-dogmatism-and-phenomenal-conservatism/.

Chudnoff, E. (forthcoming). Intuition, presentational phenomenology, and awareness of abstract objects. *Florida Philosophical Review*.

Chudnoff, E. (2016). Moral perception: High level perception or low level intuition? In T. Breyer & C. Gutland (Eds.), Phenomenology of thinking (pp. 207–220). London: Routledge.

Chudnoff, E., & DiDomenico, D. (2015). The epistemic unity of perception. *Pacific Philosophical Quarterly*, *96*(4), 535–549.

Conee, E. (2013). Seeming evidence. In C. Tucker (Ed.), *Seemings and justification: New essays on dogmatism and phenomenal conservatism* (pp. 52–70). Oxford: Oxford University Press.

Davidson, D. (1973). Radical interpretation. *Dialectica*, 27. Reprinted in *Inquiries into truth and interpretation* (pp. 125–140). Oxford: Clarendon Press, 1984.

Ferreira, F., Bailey, K. G. D., & Ferraro, V. (2002). Good-enough representations in language comprehension. *Current Directions in Psychological Science*, *11*, 11–15.

Fodor, J. (1983). *Modularity of mind*. Cambridge, MA: MIT Press.

Fricker, E. (2003). Understanding and knowledge of what is said. In A. Barber (Ed.), *Epistemology of language* (pp. 325–366). London: Oxford University Press.

Glezer, L. S., Kim, J., Rule, J., Jiang, X., & Riesenhuber, M. (2015). Adding words to the brain's visual dictionary: Novel word learning selectively sharpens orthographic representations in the VWFA. *The Journal of Neuroscience*, *35*(12), 4965–4972.

Ghijsen, H. (2015). Grounding perceptual dogmatism: What are perceptual seemings? *The Southern Journal of Philosophy*, *53*(2), 196–215.

Gilbert, C. D., & Li, W. (2013). Top-down influences on visual processing. *Nature Reviews: Neuroscience*, *14*, 351.

Gobet, F., & Simon, H. A. (1996). Templates in chess memory: A mechanism for recalling several boards. *Cognitive Psychology*, *31*, 1–40.

Gobet, F., & Simon, H. A. (2000). Five seconds or sixty? Presentation time in expert memory. *Cognitive Science*, *24*, 651–682.

Gordon, I. E. (2004). *Theories of visual perception* (3rd ed.). Hove: Psychology Press.

Longworth, G. (2008). Linguistic understanding and knowledge. *Nous*, *42*(1), 50–79.

Lotto, A., & Holt, L. L. (2011). Psychology of auditory perception. *Wiley Interdisciplinary Reviews: Cognitive Science*, *2*(5), 479–489.

Lyons, J. (2015). Seemings and justification. *Analysis Reviews*, *75*(1), 153–164.
Matlin, M. W. (2013). *Cognition* (8th ed.). Hoboken, NJ: Wiley.
McDowell, J. (1978). On 'The Reality of the Past'. In C. Hookway & P. Pettit (Eds.), *Action and interpretation* (pp. 127–244). Cambridge: CUP.
McDowell, J. (1981). Anti-realism and the epistemology of understanding. In H. Parret & J. Bouveresse (Eds.), *Meaning and understanding* (pp. 225–248). Berlin: De Gruyter.
McLaughlin, B., Rossman, T., & McLeod, B. (1983). Second language learning: An information-processing perspective. *Language Learning*, *33*(2), 135–158.
O'Callaghan, C. (2011). Against hearing meanings. *The Philosophical Quarterly*, *61*(245), 783–807.
Pettit, D. (2010). On the epistemology and psychology of speech comprehension. *The Baltic International Yearbook of Cognition, Logic and Communication*, *5*(Meaning, Understanding and Knowledge), 1–43.
Politzer, G., & Bonnefon, J.-F. (2006). Two varieties of conditionals and two kinds of defeaters help reveal two fundamental types of reasoning. *Mind and Language*, *21*(4), 484–503.
Pollock, J. (1986). *Contemporary theories of knowledge*. Savage, MD: Rowman and Littlefield.
Pollock, J. (1987). Defeasible reasons. *Cognitive Science*, *11*(4), 481–518.
Pylyshyn, Z. (1999). Is vision continuous with cognition? The case for cognitive impenetrability of visual perception. *Behavioral and Brain Sciences*, *22*, 341–423.
Recanati, F. (2004). *Literal meaning*. Cambridge: Cambridge University Press.
Reiland, I. (2015a). On experiencing meanings. *The Southern Journal of Philosophy*, *53*(4), 481–492.
Reiland, I. (2015b). Experience, seemings, and evidence. *Pacific Philosophical Quarterly*, *96*(4), 510–534.
Siegel, S. (2005). Which properties are represented in perception? In T. Szabo Gendler & J. Hawthorne (Eds.), *Perceptual experience* (pp. 481–503). Oxford: Oxford University Press.
Siegel, S. (2012). Cognitive penetrability and perceptual justification. *Noûs*, *46*(2), 201–222.
Stanley, J. (2005). Hornsby on the phenomenology of speech. *The Aristotelian Society Supplementary*, *79*, 131–146.
Suga, N., Xiao, Z., Ma, X., & Ji, W. (2002). Plasticity and corticofugal modulation for hearing in adult animals. *Neuron*, *36*(1), 9–18.
Swets, B., Desmet, T., Clifton, C, Jr., & Ferreira, F. (2008). Underspecification of syntactic ambiguities: Evidence from self-paced reading. *Memory and Cognition*, *36*, 201–216.
Tucker, C. (2010). Why open-minded people should endorse dogmatism. *Philosophical Perspectives*, *24*, 529–545.