RESEARCH ARTICLE

The Riddle of Understanding Nonsense

Krystian Bogucki*

Received: 4 May 2023 / Revised: 13 December 2023 / Accepted: 16 December 2023

Abstract: Typically, if I understand a sentence, then it expresses a proposition that I entertain. Nonsensical sentences don’t express propositions, but there are contexts in which we talk about understanding nonsensical sentences. For example, we accept various kinds of semantically defective sentences in fiction, philosophy, and everyday life. Furthermore, it is a standard assumption that if a sentence is nonsensical, then it makes no sense to say that it implies anything or is implied by other sentences. Semantically uninterpreted sentences don’t have logical characteristics. Hence, the riddle of understanding nonsense arises. We seem to use nonsensical sentences in reasoning, thinking, judging, and drawing conclusions, but they convey no propositions, which are the vehicles of their semantic properties. In this article, I propose the pretence theory of understanding nonsense to explain the riddle of understanding nonsense, and discuss alternative frameworks that are insufficient to solve it.

Keywords: Nonsense; pretence; understanding, fiction, category mistakes.

* Polish Academy of Sciences

https://orcid.org/0000-0002-4753-497X

Institute of Philosophy and Sociology, Polish Academy of Sciences, Nowy Świat 72, 00-330 Warsaw, Poland.

krystian.bogucki@ifispan.edu.pl

© The Author. Journal compilation © The Editorial Board, Organon F.

This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International Public License (CC BY-NC 4.0).
1. Introduction

The riddle of understanding nonsense is the following set of problems. Typically, if I understand a sentence $s$, then I entertain proposition $p$, which is expressed by the sentence $s$. Alternatively, if I understand a sentence $s$, then there are some truth-conditions of this sentence that I understand, or there is a mental representation of the proposition $p$ that is the content of my belief. In each case, the sentence $s$ expresses the shared content that is responsible for its understanding. Nonsensical sentences, on the other hand, don’t express propositions (or truth-conditions if you prefer this semantic framework). Nonsensical sentences have no meaning; hence, they don’t express propositions or truth-conditions. Furthermore, it is a standard assumption that if a sentence is nonsensical, then it makes no sense to say that it implies anything or is implied by other sentences (Glock 2004, White 2011, Whitersepoon 2000, cf. McManus 2014). Logical relations occur between contents of sentences: some propositions imply other propositions and are implied by another ones. Uninterpreted sentences don’t possess logical characteristics. The contents of nonsensical sentences cannot stand in logical relations since nonsensical sentences don’t have contents (or, e.g., truth-conditions).

In this article, I propose a pretence theory of understanding nonsense. A theory of understanding nonsense is a theory that aims to explain how we understand nonsensical sentences, how we can draw consequences from nonsensical sentences, how we can reason with nonsensical sentences, and how our understanding of nonsensical sentences differs from our understanding of meaningful sentences. In short, the pretence theory of understanding nonsense is supposed to indicate the mechanism behind understanding nonsensical sentences, and it is a solution to the riddle of understanding nonsense. The concept of pretence has been used to explain a wide range of phenomena: there are pretence accounts of fiction (e.g., Walton 2000).

---

1 The details of these stories depend on the preferred theory of language: whether it should be expressed in terms of propositions, truth-conditions, the language of thought, etc. The story about what meaningful sentences express can also be formulated in terms of use. These reservations will be taken into account in the later parts of the article.
1990, 2015), existence (e.g., Evans 1982, Walton 1990), truth and reference (e.g., Armour-Garb and Woodbridge 2013, 2015), mathematics (e.g., Armour-Garb and Woodbridge 2015, Yablo 2001), semantics of attitude ascriptions (Crimmins 1998) and many others. The notion of pretence is also used to account for some specific features of other kinds of discourse; for instance, most theories of fiction appeal to it in some way (e.g., Lewis 1978, Searle 1975, Thomasson 1999, 2003). So, it has already been proved that the notion of pretence is fruitful, and I believe that the notion of pretence is the correct solution to the riddle of understanding nonsense.

The paper is organised as follows. In Section 2, I will show why the riddle of understanding nonsense is important. In Section 3, I will present a framework that is the correct answer to the riddle of understanding nonsense. In Section 4, I will point out the shortcomings of Manish Oza’s (2022) pretence account of nonsense. In Section 5, I will show that L. J. Keller and J. A. Keller’s (2021) and Recanati’s (1997) positions, which attempt to solve the riddle of understanding nonsense by means the language of thought hypothesis, are not satisfactory.

2. The significance of understanding nonsense

In my view, there are some important domains in which the notion of understanding nonsense is useful and widespread. The first such domain is fiction of all kinds, especially fairy tales and children’s literature. Fairy
tales commonly describe magic events and unusual characters; many of their sentences involve category mistakes. In a fairy tale, for instance, a pot can talk, see and hear. It can drink coffee and talk to other characters. There are no problems with these actions in the imagined world of fiction; however, from the point of view of the most prominent theory of nonsense, sentences describing these actions are problematic. Let’s note the sentence

(1) The pot is drinking coffee.

This sentence is nonsensical because it attributes to an inanimate thing a property that can only be ascribed to other types of stuff (animate objects). Only living things can breathe, eat and drink. Of inanimate objects, it might make sense to say that they are heavy, pretty or ceramic. Structurally (1) is similar to (2) and (3):

(2) Julius Caesar is a prime number.

that recent work on category mistakes has made people more critical of what I call below the standard view, which postulates that category mistakes are nonsensical (Camp 2004, Magidor 2009, 2013).

5 The term “category mistakes” is ambiguous. On the one hand, it can denote a class of sentences that are semantically or pragmatically infelicitous. This way of speaking (let’s call it “semantical”) just identifies an appropriate type of sentence in order to pose the problem of the nature of its infelicity. Ofra Magidor, for instance, uses the term “category mistakes” in this way in the title of her paper “Category Mistakes are Meaningful”. In her book, “Category Mistakes”, she does not define the titular notion but introduces it through examples (2013, 1). This way of speaking (let’s call it “logical”) stipulates that category mistakes are nonsensical because of a violation of logical syntax. I use the term “category mistake” in the former sense. This way of using the term “category mistake” is neutral towards different theories of category mistakes. This is a very important feature since it makes it possible to talk about sentences like “Julius Caesar is a prime number” without siding with any conception of nonsense. Proponents of the no nonsense view, such as Magidor, claim that category mistakes are semantically meaningful (though pragmatically infelicitous). Adherents of the austere approach claim that category mistakes are nonsensical, but their account of the source of nonsensicality differs from the standard view of nonsense.

6 This example has been discussed by Ludwig Wittgenstein in Philosophical Investigations (§281). For a discussion, see Glock (2004, 241).
(3) The theory of relativity is eating breakfast.

Each of these sentences contains some sort of category mistake, and according to the standard view of nonsense these category mistakes are violations of the rules of logical syntax. The standard view of nonsense holds that the main source of nonsense is sentences that violate the rules of logical syntax. Some sentences violate the rules of logical syntax because their components cannot be connected on the basis of logical categories of words. The standard view postulates the division of words into logical categories (e.g., material objects, properties of material objects, properties of properties of material objects, etc.; abstract objects, their properties, properties of these properties, animate objects and their properties, etc.). Sentences containing words with incompatible logical categories of this sort are nonsensical. (1), (2) and (3) violate the rules of logical syntax and are nonsensical. If one asserts (2), one makes a category mistake of attributing the property of being an abstract object to a material object. If one asserts (3), one makes a category mistake of ascribing to an abstract object a property enjoyed only by animate objects. In fact, sentence (2) is Rudolf Carnap’s famous example illustrating how sentences that violate the rules of logical syntax result in category mistakes and thus in nonsensical statements (Carnap, 1959). Carnap says in The Elimination of Metaphysics Through Logical Analysis of Language:

---

7 The most prominent defenders of the standard view are Russell (1908, 1910), Carnap (1959) and Ryle (1938, 1949). More recent defences of it can be found in Hacker (2003) and Glock (2015). Some Wittgensteinians call this approach to nonsense “the substantial view of nonsense” because it distinguishes between “mere” nonsense (e.g., “John is xwwwy”), which contains a component without meaning, and a more substantial kind of nonsense (e.g., “Julius Caesar is a prime number”), which requires a notion of logical syntax. Carnap (1959) and Hacker (2003) also give a fairly clear description of what logical syntax is.

8 Ryle (2009, 178) characterises category mistakes as follows:
When a sentence is (not true or false but) nonsensical or absurd, although its vocabulary is conventional and its grammatical construction is regular, we say that it is absurd because at least one ingredient expression in it is not of the right type to be coupled or to be coupled in that way with the other ingredient expression or expressions in it. Such sentences, we may say, commit type-trespasses or break type-rules.
Another very frequent violation of logical syntax is the so-called “type confusion” of concepts. (...) An artificial example is the sentence we discussed earlier: “Caesar is a prime number”. Names of persons and names of numbers belong to different logical types, and so do accordingly predicates of persons (e.g., “general”) and predicates of numbers (“prime number”) (Carnap 1959, 75).

Fiction offers much more nonsense than the aforementioned category mistakes. It is common in Wittgensteinian literature to discuss various uses of nonsense in Lewis Carroll’s Alice in Wonderland books (Glock 2004, 2015, White 2011, McManus 2014). These literary examples include nonsense poems such as Jabberwocky, nonsensical uses of the word ‘nobody’, and metaphysically impossible events, such as the appearance of the Cheshire Cat. We seem to understand what is going on in all these passages, even though they violate the rules of language. Jabberwocky consists of nonsensical words invented by Carroll. The uses of the word ‘nobody’ violate the standard syntactic role of ‘nobody’. The word ‘nobody’ is used as if it were a proper name rather than an indefinite pronoun. The case of the Cheshire Cat is slightly different as it belongs to the same group as Mauritius C. Escher’s works of art (Hacker 2003, White 2011, McManus 2014), which present metaphysically, logically, or mathematically impossible events and produce visual illusions in the viewer.9 Furthermore, some other techniques used in fiction, such as tropes of personification (e.g., abstract entities like death can ride horses or have hands) or metafictional techniques (e.g., fictitious protagonists can address the audience or ‘chase’ the author of the book), also constitute exemplary cases of category mistakes (Nolan 2021). Again, according to the standard view of nonsense, category mistakes are primary examples of nonsense since they violate the rules of logical syntax. Yet, it is plausible that we somehow understand these nonsensical stories.

9 One can question whether these are real instances of nonsense. At least some Wittgenstein scholars would vehemently argue that they are. For the record, Glock (2004, 239) and White (2011) provide other literary and non-literary cases of understanding nonsense that seem far more controversial to me. However, this shows that the class of potential candidates is quite wide.
Every reader of *Tractatus Logico-Philosophicus* should address the issue of understanding nonsense. Wittgenstein famously claims that many/all sentences (*Sätze*) in his works are nonsensical (*unsinnig*) (§6.54):

> My sentences [*Sätze*] serve as elucidations in the following way: anyone who understands me eventually recognizes them as nonsensical [*unsinnig*], when he has used them— as steps— to climb up beyond them. (He must, so to speak, throw away the ladder after he has climbed up it.)

> He must overcome [*überwinden*] these propositions, and then he will see the world aright.

Wittgenstein also gives some specific examples of sentences that he says are nonsensical, such as ‘There are objects’ and ‘1 is a number’ (§4.1272). However, we can set aside these two controversial instances and focus on another Tractarian example:

\[(4)\quad \text{A picture is a fact.}\]

In fact, Wittgenstein claims that it is not possible to say anything about formal concepts such as complexes, facts, functions, numbers and pictures without falling into nonsense. *Tractatus* is full of statements of this kind (§4.1272). It very often says things like “a fact is such-and-such”, “a function is such-and-such”. The conclusion that the *Tractatus* is nonsensical is justified in its own terms. For this reason, the riddle of understanding nonsense arises; the reader thinks she understands Tractarian sentences and draw some conclusions from Tractarian arguments. If Tractarian sentences are nonsensical, then there are no propositions that the reader can understand. If no propositions are expressed by Tractarian sentences, then there are no logical inferences that the reader can entertain. Nonsensical sentences do not possess the logical characteristics which meaningful sentences enjoy, but, at first sight, Tractarian sentences aren’t psychologically distinguishable from the latter. It seems inexplicable why we seem to understand Tractarian sentences, why we read them like ordinary, meaningful sentences, and how they can exert any influence on readers. Furthermore, as noted by Peter Sullivan (2003), the influence of *Tractatus* should be conceptual, not causal: it should persuade by means of concepts. The influence on readers of *Tractatus* is not ‘like a blow on the head’
The engagement and appreciation of Tractarian nonsense is conceptual in nature.

The issue of understanding Tractarian sentences may seem to be just a historical problem or a curiosity caused by Wittgenstein’s excessively inflated theory of syntax. After all, it takes some time to understand why Tractarian sentences are nonsensical. However, it would be good to have a theory explaining such a cardinal problem in an exegesis of a twentieth-century masterpiece. Furthermore, Carnap had a tendency to accuse most philosophical classics of talking nonsense. Some of these accusations were based not on violations of the principle of verification, but on violations of logical syntax. As such, they are still valid to some extent, and, as we shall

---

10 In this paper I can only suggest pretence as the mechanism behind understanding Tractarian nonsense. I’ll devote another paper to the details of this view. My account is broadly in agreement with Cora Diamond’s remarks on understanding Tractarian nonsense through imagination (2000, 157–160), but I would like to reinterpret imagination in terms of pretence. In her paper, Diamond observes that a nonsense sentence expresses no proposition, yet “to understand a person who utters nonsense is to go as far as one can with the idea that there is [a proposition]” (Diamond 2000, 157). I share with Diamond the framework of the austere view and a resolute reading of Wittgenstein’s early thought, and I agree with some of her specific remarks on the topic of understanding a person who talks nonsense. For instance, she rightly recognises that we are neither inside nor outside that person’s thought. At the same time, I appreciate McManus’s criticism of Diamond’s remarks on imagination: Diamond’s suggestion is no more than a proposal that needs an elaboration that it has not yet received (McManus 2014, 171). It’s not clear how Diamond reconciles the claim that nonsensical sentences have no content with her specific claims about the status of ethical and philosophical theses. How can Tractarian sentences exert any influence on the reader if these sentences convey no propositions? As far as I can see, Diamond’s proposal amounts to a correct suggestion of a proposition-like structure of the imagination (Diamond 2000, 157, the words quoted above), and the false idea that attributions of nonsense are themselves nonsensical (Diamond 2000, 157–158). (Of course, Diamond’s paper offers some fundamental views on the ethical dimension of the Tractatus and a penetrating critique of nonsensical sentences that “have something, something true, but unsayable”). I think that neither Diamond’s (2010) remarks on the transitional sense of philosophical remarks nor Conant’s views on the clarification of thoughts (2001, 60) are much more helpful in understanding nonsense than her remarks on “imagination”.

*Organon F* 30 (4) 2023: 372–411
see in the next paragraph, other concerns have recently been raised about the validity of much philosophical discourse. It would be good to know the mechanism of understanding philosophical works that are allegedly nonsensical.

A third domain of nonsense has recently been indicated by Herman Cappelen (2012, 2013). We, philosophers, like to think that our own assertions and statements are meaningful, and nonsense is a problem of the dead. Some historical philosophers made irresponsible claims and were irrevocably unclear about what their thoughts and claims meant; however, this does not happen now, in the twenty-first century (!), because the main characteristic of analytic philosophy is its clarity. Cappelen’s Verbal Virus Theory (2012, 49–60) challenges the view that contemporary analytic philosophy is free of nonsense. He argues that the term ‘intuition’ and its cognates are semantically defective because they fail to have content, thus leading to sentential nonsense, assertoric nonsense, and cognitive nonsense. 11 Hence, for example, this sentence (5) is semantically defective:

(5)  Intuitively, the Twin Earth liquid is not water.

‘Intuition’ is one of the most widely used philosophical terms of the second half of the twentieth century, but how did a meaningless term come to infect so much work? Here comes the virus part of the theory: as stated by Cappelen (2012, 50), ‘Philosophers’ use of “intuition” is a kind of intellectual/verbal virus (or tick) that started spreading about thirty to forty years ago’. What is the source of this virus? There are some clues, but this issue requires more historical analysis (Cappelen 2012, 56–57, Hintikka 1999).

Cappelen doesn’t limit himself to claiming that only the term ‘intuition’ has such a miserable fate. No, its fate is far more common in contemporary philosophy. Cappelen postulates that the terms ‘semantics’, ‘pragmatics’, ‘a priori’, ‘a posteriori’, ‘justification’, ‘causation’, ‘evidence’ and ‘person’ also fall in the category of semantic defectiveness (Cappelen 2012, 60, fn.

11 I borrowed this classification from (L. J. Keller and J. A. Keller 2021). X is a sentential nonsense if and only if x is a sentence that lacks content. Y is assertoric nonsense if and only if y is an assertion that lacks content. Z is cognitive nonsense if and only if z is a thought (belief, hope, desire, etc.) that lacks content. See Cappelen (2013, 26) for his own division of defective types of content.
10. So, in his view, nonsense is quite widespread in contemporary philosophy. There are several reasons why these concepts are flawed. Roughly speaking, each of these terms is a theoretical term that has been defectively introduced. A term $T$ is defectively introduced if it has one of the following characteristics (Cappelen 2012, 52):

- $T$ has no agreed-upon definition among practitioners of a discipline.
- There is no agreement among practitioners of a discipline regarding which cases constitute core paradigms of the extension of $T$.
- There is no agreed-upon theoretical role that $T$ (or $T$'s extension) plays in a discipline.
- There is considerable disagreement and dispute within a discipline about $T$ itself.

This is only meant to be a rough outline, and it was later reformulated (Cappelen 2013, 38–40). However, it gives the reader an idea of what can be problematic about these terms. It’s not surprising that terms such as ‘semantics’ or ‘a priori’ are used differently by scholars, and the discussion about what they mean is ongoing. It’s not obvious how such terms acquire meaning in philosophical discourse, because experts disagree about their content. One possibility is that they inherit meaning from their natural language counterparts (Cappelen 2012, 27). I think that Cappelen’s view is plausible for those terms that don’t have a clear usage in natural languages. Such terms are ‘intuition’, ‘semantics’, ‘pragmatics’, ‘a priori’ and ‘a posteriori’. They don’t have fixed use in ordinary language and they seem to satisfy Cappelen’s diagnostics.13

Cappelen (2013, 36) doesn’t think that philosophers are particularly prone to nonsense, and he describes the practice of philosophers criticising

---

12 Van Inwagen (1980) claims that philosophers’ use of the word “body” is nonsensical, since there is no correct definition of this term. Furthermore, he suggests that such a definition isn’t possible, and philosophers should give up using this term to avoid talking nonsense (cf. Tye 1980).

13 Cappelen (2012, 29-48) argues convincingly that ‘intuition’ and its cognates don’t have a common and uniform meaning in English. They can refer to a variety of states, events and things, and their meaning is context-dependent.
other philosophers for talking nonsense as somewhat objectionable. Be that as it may, he thinks that nonsense is generally more common than we think, and there are some good reasons for this. Speakers are fallible in respect to grounding facts of content. It may turn out that a term we use was introduced in a defective way, or that it was defectively transmitted through a chain of communication. 14 Cases of defective introduction include a lack of appropriate naming, demonstration, or intention on the part of the person introducing a term. I think language usage abounds with such cases. One such case has recently been discussed in the literature on demonstrative thoughts and various illusions of thought (L. J. Keller and J. A. Keller 2021, O’Brien 2009, Stojnić and LePore 2020). On the other hand, even if defective transmissions aren’t widespread, they can happen. If communication is effective, it is so because it is carried out carefully. This condition isn’t always met, and semantic failures occur.

To sum up, there are three distinct areas of potential philosophical investigation: I) fictional discourse; II) historical examination of Wittgenstein’s, Carnap’s, and others’ views; III) illusions of thought as well as more contemporary discussions of notions such as “intuition”, “semantics”, “pragmatics”, etc. I have chosen to group these areas in this way because all these groups are heterogeneous and propositions belonging to them may be semantically defective for different reasons: group I) consists of, among other things, category mistakes, poems made up of neologisms, and sentences that describe impossible states of affairs; group II) isn’t uniform either. According to some commentators, nonsensical Tractarian sentences are violations of logical syntax (Glock 2004, Hacker 2003). Other scholars deny this and claim that they are just mere nonsenses (Conant 2003, Diamond 1978, 1981). Thus, depending on our interpretation of Wittgenstein, we may see his relation to Carnap’s position on nonsense, which focused on violations of logical syntax, differently. The works of Hegel, Descartes and other philosophers may contain nonsense of other kinds. Van Inwagen (1980, 285), for example, claims that Descartes’ Meditations on First

---

14 The distinction between failures of introduction and failures of transmission is presumably not disjoint. I can unsuccessfully name someone and then pass this name to someone who, in turn, will transmit it ‘incorrectly’ (e.g., because she mishears me). J. J. Keller and J. A. Keller (2021) consider some possible scenarios of failures.
Philosophy contains nonsensical passages. Van Inwagen’s work may suggest that these nonsensees are of the Cappelenian type, i.e., they result from a lack of agreement among experts on the correct definition of a term, or from a failure to introduce a term correctly. III) covers cases of a different kind from the Carnapian type. They have nothing to do with logical syntax and are closely related to the introduction of theoretical concepts. Cappelen believes that philosophy may have an important, purely cognitive function, but he claims that some terms in philosophy have not been properly introduced and thus result in nonsense. Nevertheless, other disciplines may produce nonsense of the same kind. In contrast to Cappelen, both Wittgenstein and Carnap were convinced that philosophical discourse is inherently flawed and philosophical errors are of a radically different kind from those in other domains.

I believe that this section has shown that there are some contexts in which we seem to accept various kinds of semantically defective sentences. We engage in the practice of asserting and rejecting various sorts of nonsensical sentences in fiction, philosophy, and everyday life, thus it is hard to deny that nonsensees appear in the business of reasoning, thinking, and judging. A pretence theory of understanding nonsense will explain how it is possible to engage in practices involving the use of nonsensical sentences in the specified domains. In the next section, I will describe the view which solves the riddle of understanding nonsense; also, I will present the rules of the game of make-believe that are responsible for the mechanism of understanding nonsense.

3. A pretence theory of understanding nonsense

I begin the presentation of my account with a few observations. First, I agree with the standard assumption that nonsensical sentences express neither propositions nor truth-conditions. To have meaning is to express some content, so nonsensicality must amount to a lack of content. This assumption is sometimes rejected (Sorensen 2002). I also accept the view that nonsensical sentences do not stand in logical relations. Only propositions imply other propositions or are entailed by other propositions. This view is also sometimes rejected (McManus 2014). Furthermore, I deny that nonsensical
sentences have logical forms (or partially interpreted logical forms). Oza (2022) claims that logical forms characterise sentences in general, and he argues that this view is necessary to explain the mechanism of understanding nonsense.\(^\text{15}\) In my opinion, uninterpreted sentences are just physical objects and don’t have logical properties. Sentences acquire logical forms if and only if they have content: they express propositions or truth-conditions, \textit{etc.} Contents have logical forms, and we can derivatively ascribe logical forms to sentences. Sometimes, two or more sentences express the same content. In such cases, different propositions have the same logical form. The account that only propositions have logical forms seems to me no less standard than the first two assumptions.\(^\text{16}\) For example, if one accepts Donnellan’s (1966) view of the semantics of referential and attributive uses,

\(^{15}\) There is an additional complication in Oza’s (2022) views on the relation between pretence and logical form. According to him, we pretend that a nonsensical proposition expresses a logical form, but the pretended logical form is the one that the nonsensical sentence really has. This is not as unusual as it sounds: in fact, it is quite common for us to pretend to have some properties that we really have (see Langland-Hassan 2014, 11). For example, if I were to pretend that I am Wittgenstein, I would have to pretend that I am an Austrian, a philosopher, a human being, that I have a body and two legs, and that I have lived in Vienna for some time. But in fact I am a philosopher, a human being, I have a body and two legs, and I have lived for some time in Vienna. In contrast to Oza’s position, my theory holds that we make believe that a nonsensical sentence expresses a logical form that is merely suggested by the nonsensical nonsense. Oza’s stance helps him to show the difference between his view of nonsense and Diamond’s position. He also believes that we can understand the logical form of a sentence that we don’t understand (Oza 2022, 16–17). I disagree with this. I understand neither the meaning nor the logical form of the sentence “Postupně se vynořoval ucelený obraz genealogických souvislostí” because I don’t speak Czech (I found it on Jaroslav Peregrin’s website). Oza’s view is well summarised in this fragment (Oza 2022, 20, fn. 29): “Given the role of form in my account, a string which is totally empty of form – say ‘xg7*12d’ – will, without further set-up, not be the subject of a pretence. Thus, ‘it is nonsense that xg7*12d’ can only be read metalinguistically. I think this is the correct result. To the extent that a nonsense sentence has some (even partial) syntactic form, there is the possibility of a pretence and a non-metalinguistic reading of the nonsense-attribute”.

\(^{16}\) I think it is shared by Higginbotham (1993) and LePore and Ludwig (2002), to whom Oza refers in other parts of his paper. In the context of logical structure of
then one must admit that the logical form of the sentence ‘Smith’s murderer is insane’ depends on whether our intention was to express a truth-condition involving a quantifier expression or a truth-condition involving a singular term. Clearly the proposition remains the same, but its logical form is different. Moreover, if one accepts a wide range of contextual dependencies, then one has to acknowledge a greater degree of independence between a sentence and its logical form (Travis 1994, 2017). At present, contextualism seems to dominate in the dispute over the scope of contextual dependence.

Nonsensical sentences don’t express propositions; they don’t have a logical structure and they don’t stand in logical relations. So, what is the mechanism behind understanding nonsensical sentences? It’s pretence. When a speaker communicates using the nonsensical sentence $s_1$, she expresses the make-believe proposition $p_1$ within a game of make-believe. A hearer understands her correctly if and only if she grasps the make-believe proposition expressed by the speaker. More precisely, the hearer understands the nonsensical sentence $s_1$ if and only if she grasps the make-believe proposition $p_1$. There is a game of make-believe that specifies the proposition expressed by the nonsensical sentence.$^{17}$ Furthermore, when an agent derives the nonsensical sentence $s_2$ from the nonsensical sentence $s_1$, she pretends that the proposition $p_1$, which is expressed by the sentence $s_1$ in our game of make-believe, implies the proposition $p_2$, which is expressed by the sentence $s_2$ in our game of make-believe. The nonsensical sentence $s_1$ implies the sentence $s_2$ if and only if there is the pretended proposition $p_2$, which is entailed by the pretended proposition $p_1$ (in our game of make-believe).$^{18}$ A logical entailment is successfully communicated if and only if

$^{17}$ In a sense that will be explained in a moment. For now, let’s stick to a more informal way of speaking.

$^{18}$ There is a multitude of games of make-believe. Every inference and truth within pretence is relative to a game of make-believe. I think that works of fiction (books, plays, paintings) are games of make-believe, but for the purposes of this paper this sentence, he refers to the book by Armour-Garb and Woodbridge (2015), but these authors explicitly say that a hearer understands a logical form of some semantically defective sentences (e.g., a Liar sentence), but not full-blown nonsense such as ‘colourless green ideas sleep furiously’ (Armour-Garb and Woodbridge 2013, 849, fn. 6, 2015, 158, fn. 13). Thus, Armour-Garb and Woodbridge’s view differs from Oza’s position.
the hearer grasps the logical entailment make-believably expressed by the speaker. Finally, a speaker pretends that the proposition \(p_1\), which is make-believably expressed by the nonsensical sentence \(s_1\), is of the logical form \(LF_1\); to grasp the make-believe proposition correctly, the hearer should correctly recognise the make-believe logical form.

Propositions in games of make-believe don’t inherit logical structure from nonsensical sentences. Logical structure is another element that is pretended in a game of make-believe. The reason for this is well known: sentences (especially nonsensical ones) don’t have a logical form. A make-believe proposition doesn’t have to reflect the logical form of a sentence because nonsensical sentences can have incoherent logical forms that are not instantiated in the world. If the rules of logical syntax are supposed to specify prohibited combinations of logical categories (Hacker 2003, 7), then prohibited combinations should occur neither in propositions nor, hence, in language, which is a set of contents of sentences and speech acts. Category mistakes lead to nonsensical sentences that don’t express propositions and have no logical structure. Logical forms which are forbidden by the rules of logical syntax do not occur as such. Let’s take sentence (6) as an example:

(6) Chairman Mao is rare.

It is claimed that (6) is nonsensical (Dummett 1983, 51). ‘Chairman Mao’ is a name-like expression, and it requires a first-order predicate in order to be meaningfully combined (e.g., is bold, is a man). ‘Rare’ is a second-order predicate, and it can be ascribed to first-order predicates (e.g., an honest politician, gold).\(^{19}\) (6) is supposed to combine two incompatible logical forms: ‘\(a \, \varphi\)’ and ‘\(\psi \, \chi\)’.\(^{20}\) (6) is a nonsensical sentence and doesn’t express any proposition. In particular, it doesn’t express the proposition of the logical form that combines ‘\(a \, \varphi\)’ and ‘\(\psi \, \chi\)’. So, there is no proposition of this logical form. There are only some sentences which have a surface grammar in a form that combines ‘\(a \, \varphi\)’ and ‘\(\psi \, \chi\)’. In my view,

\(^{19}\) See Conant’s discussion of this example (2002, 395–398, 403–405).
\(^{20}\) Lowercase Greek letters stand for first-order concepts, while uppercase Greek letters stand for second-order concepts.
the logical structures that are pretended in a game of make-believe are based on the surface grammar of sentences. The pretended proposition expressed by a nonsensical sentence usually has the logical form that it suggests. It’s not the same logical structure because nonsensical sentences have no logical structure. Furthermore, it’s a regularity that the surface grammar of a sentence is often the same as the logical form of the proposition, but there are exceptions. If someone were to express the nonsensical sentence ‘The King of England is a prime number’, then the proposition expressed in a game of make-believe would involve a quantifier-like statement. The surface grammar of the sentence is different from the logical form of the expressed proposition.

When it comes to understanding nonsense, pretence functions as a stable mechanism that is exploited by a language user. The user simply exploits the appropriate convention of make-believe to achieve her communicative aims in accordance with the existing norms. She does not need to intend to pretend, nor does she need to be aware that she is pretending. The opposite appearance may stem from the ordinary meaning of the term “pretence”. However, intuitions that link pretence with an intentional or deliberate act are incorrect (Armour-Garb and Woodbridge 2015, 63; Walton 1990, 38; Yablo 2001, 97). A pretence theory is concerned not with the mental states and introspection of speakers but with the linguistic functioning of some parts of discourse. By analogy, a theorist of fiction does not claim that an artist or her audience are aware of any pretence going on, and even less does a mathematician think that pretence is necessary to make any claims in his beloved field of inquiry. Walton even analyses dreams (1990, 43–50) and metaphors (1993) as types of pretence. So, what does the pretence episode look like from the speaker’s point of view? Her knowledge is dispositional and practical. She proceeds, as Armour-Garb and Woodbridge (2015, 71) suggest, from the general assumption that names and designation expressions refer to objects and that predicates describe objects. She does not feel compelled to stop speaking in certain ways when confronted with putative obstacles to her practice, and she does not take a stance on how her sentences work. She simply uses these propositions as a tool to make

Assuming the attributive reading of the description.
assertions. Furthermore, the speaker is not required to keep the rules of make-believe consciously in mind. In general, the rules of make-believe are not explicitly agreed or formulated, and the participants in a game of make-believe may not be aware of them. However, they are internalised by speakers: the rules of make-believe prompt the actions of agents, and in this sense they are operative in practice because speakers act upon them. For this reason, we have to observe the rules of make-believe in the practice of speaking a language in order to formulate them explicitly. Nevertheless, from the theoretical point of view, we should consider that the speakers are engaged in a game of make-believe and understand these fragments of discourse as involving pretence.22

There is a standard conception of games of make-believe, but the details of how to understand them vary (Armour-Garb 2015, Crimmins 1998, Evans 1982, Oza 2022, Walton 1990, 2015). The rules of make-believe can be divided into direct and indirect ones. Direct rules specify outright the basic principles of the game of make-believe; they include props and other stipulated pretences. Indirect rules prescribe what is make-believe based on the direct rules and real-world conditions, and they can be divided into principles of generation and recursive principles. Thus, the rules of a game of make-believe typically include:

a) Props,
b) Direct principles,
c) Indirect principles: Principles of generation (PG) and Recursive principles (RP).

a) and b) establish pretences that are expressly make-believe. For instance, in a children’s game of make-believe they establish that a lump of mud is a biscuit and that some gestures towards the mouth are an act of eating the biscuit. Thanks to c), participants of a game of make-believe can draw upon their general knowledge of the world to generate the propositions, facts and events prescribed by the game. (PG) and (RP) aim to make the make-believe world as similar to the real world as the direct rules allow. For

example, in the absence of any indication to the contrary, we assume that all the protagonists of a novel eat, sleep, have blood in their veins, have ambitions, dreams, and so on. We assume that the fictional world is quite similar to our own. Of course, there are some differences between pretense episodes and their real-world counterparts. In sum, games of make-believe set out local conventions that govern how some features of the actual world are to be mapped to some make-believe circumstances.

The proposed theory of understanding nonsense is semantic. The props for this game of make-believe are all linguistic expressions (unlike, for example, children’s make-believe games, but similar to the rules of fiction). The rules are as follows:

**The Game of Make-Believe of Understanding Nonsense**

**A. Props:**

The props for the game are the following types of linguistic expressions: names, designation expressions (including demonstratives and definite descriptions), predicative expressions (ascribing properties, relations *etc.*), and other types of linguistic expressions (quantifiers, logical connectives, *etc.*).

The following pretences are prescribed for these types of expressions.

(A.1) Every name or designation expression has a bearer.

(A.2) Every predicative term expresses a property or a relation.

(A.3) Every other expression has a semantic value.

---

23 See Gendler (2003) on this topic. The differences have an impact on the way viewers receive fiction, especially the way they are emotionally engaged in the cinema (Gendler 2000, Liao and Gendler 2020, supplement: *Puzzles and Paradoxes of Imagination and the Arts*).

24 Oza (2022, 18) proposed a set of pretences for understanding nonsense. However, the rules formulated in his work are not general enough to generate pretences for different cases of understanding nonsensical sentences. Oza’s rules provide no more than an illustrative description of the pretences of the sentence ‘Goodness is hexagonal’.
B. Direct principles:

(B.1) It is to be pretended that predicative expressions describe objects (denoted by names) as having or lacking some properties.

(B.2) It is to be pretended that sentences formed from props express propositions.

(B.3) It is to be pretended that sentences formed from props have logical form.

(B.4) It is to be pretended that sentences formed from props stand in logical relations.

(B.5) The pretences displayed in sentences formed from props are prescribed if and only if they have been asserted or used in other speech acts.

C. Indirect principles:

The principle of generation:

If \( P \) is true, and if there is no set of make-believedly true sentences \( Q_1 \ldots Q_n \) such that if \( Q_1 \ldots Q_n \) were true then \( P \) would not be true, then \( P \) is make-believedly true.

Recursive Principle:

If \( P_1 \ldots P_n \) is a set of make-believe truths, and the counterfactual ‘If \( P_1 \ldots P_n \) were true, then \( R \) would be true’ is true, and there is no set of make-believe truths \( Q_1 \ldots Q_n \) such that the counterfactual ‘\( Q_1 \ldots Q_n \) were true, then \( R \) would not be true’ is true, then \( R \) is make-believedly true.\(^{25}\)

(B.5) sets the condition for engaging in a game of make-believe of understanding nonsense. It is quite liberal, but it indicates that, in general, not much is required of a language user in order to engage in the prescribed make-believe. It is enough to read sentences from a novel to engage in the pretence prescribed by it. If there is an appropriate set of conventions, then

\(^{25}\) How to correctly formulate \((PG)\) and \((RP)\) is, to some extent, an open question. It may depend on the subject of a game of make-believe. See the discussion in Walton (1990, 144–161).
the mere fact of using props (e.g., words, sentences, etc.) is enough to engage in some discourse. Furthermore, Evans (1982, 256) notes that (PG) and (RP) are not sufficient to introduce make-believe propositional attitudes. It is generally not true that if one believes the truth of ‘The pot is drinking coffee at $t_1$’, then one believes the truth of ‘the pot is not thirsty at $t_2$’. People tend to believe that pots don’t drink coffee and pots aren’t thirsty. Or, strictly speaking, people tend to believe that the sentence ‘The pot is drinking coffee at $t$’ is nonsensical and the sentence ‘The pot is not thirsty’ is nonsensical. Evans (1982, 257) proposes two principles for incorporating (some) propositional attitudes into a game of make-believe ($^*()^*$ stands for ‘it is make-believedly the case that’):

$$(x) \text{ (If } x \text{ believes that } ^*P^*, \text{ then } ^*x^* \text{ believes that } P^*)$$

$$(x) \text{ (If } x \text{ intends that } ^*P^*, \text{ then } ^*x^* \text{ intends that } P^*).$$

The general rules of the game of pretending to understand nonsense prescribe pretences for particular instances of nonsensical sentences. For example, the rules prescribed for (1) are:

**Props:**

The props are the terms ‘the pot’, ‘is’, ‘drink’, and ‘coffee’.

**Direct principles:**

The sentence ‘The pot is drinking coffee at $t_1$’ expresses the proposition that the pot is drinking coffee at $t_1$. The make-believe proposition has the logical form of ‘$a$ is $\phi$-ing at $t_1$’.

**Indirect principles:**

The truth ‘The pot is drinking coffee at $t_1$’ entails the truth ‘the pot is not thirsty at $t_2$’.

---

26 See Evans’ (1982, Ch. 10) discussion of perceptual illusions. See also Armour-Garb and Woodbridge’s (2013, 846, fn. 1, 2005, 67) rather general formulations of (what I take to be) the condition for engaging in some kind of pretense. I have benefited from their discussion of the rules of games of make-believe in these works.

*Organon* F 30 (4) 2023: 372–411
Some utterances of nonsensical sentences seem to be semantically defective from the very beginning (e.g., “Julius Caesar is a prime number”), while others don’t (e.g., “A picture is a fact”). What they have in common is that, taken literally, they do not express a proposition. The main reason why a nonsensical sentence is defective is the semantic fact that it expresses no content. However, the fact that a sentence expresses no content may be unknown to the speaker who has decided to use it. On the other hand, if we follow the literal reading of nonsensical sentences, then any use of them would turn out to be semantically infelicitous. An appeal to semantic redirection involving pretence helps us to avoid ending up with semantically infelicitous assertions. Thanks to this appeal to pretence, we can explain how a semantically defective fragment of language serves any serious purpose at all, thus solving some philosophical puzzles by recognising pretence at work. In other words, the nonsensical sentence $s_1$ does not express a proposition, so to use this sentence to assert something we have to find another way that would be governed by rules of make-believe and real-world conditions. What makes an act of pretence appropriate as a move in the game of make-believe are direct principles, indirect principles, and some assertions made by the participants. Furthermore, pretence is intrinsic to understanding nonsense in the sense that the typical linguistic functioning of nonsensical sentences involves an appeal to it because pretence-free uses lead to assertoric nonsenses.

The rules of make-believe and real-world conditions generate some acceptability conditions associated with an utterance of the nonsensical sentence $s_1$. These acceptability conditions make an utterance of $s_1$

---

27 I use the word “literally” only for lack of a better word. Some of the connotations it can carry are problematic. For example, literal should not be taken here in direct opposition to metaphorical discourse. Armour-Garb and Woodbridge (2015, p. 6, fn. 21) use the phrase “at face value” instead of “literally”. From my perspective, “at face value” can be even more problematic because surface readings of nonsensical sentences are diverse.

28 In this section, I have made use of some of the terminology introduced by Armour-Garb and Woodbridge (2015, chapters 1 and 2).

29 In his otherwise critical article on the theory of pretence, Mark Richard (2013, 191) notes that the pretence account is not an account of what is said by the uses...
appropriate or inappropriate; however, this sentence is still semantically
defective (nonsensical) because it does not express a proposition. In practi-
cal terms, this means that if we use nonsensical propositions in our premises,
then we must (in the end, in a serious mode) discard any conclusions based
on the assumption that they are true. On the other hand, we can talk about
an understanding of nonsensical sentences because the rejection of nonsen-
sical sentences and of reasoning with nonsensical sentences is parasitic on
pretence: it exploits the props, the direct principles, the indirect principles
and the real-world conditions. When we reject nonsensical propositions, we
exploit not only the fact that within a game of pretence there is a clear
difference between being able to follow inferential relations and not being
able to follow them, but also the fact that within a game of pretence there
is a clear difference between the appropriateness of some propositions and
the inappropriateness of other propositions.

In the next two sections, I will describe rival theories of understanding
nonsense and show why they are unsatisfactory. Section 4 discusses Oza’s
(2022) pretence account of nonsense and explains why my pretence theory
of understanding nonsense is more promising. Section 5 examines how L. J.
Keller and J. A. Keller (2021) interpret the language of thought to make
room for understanding nonsense.

4. Logical form and the characterisation of nonsense

I would like to begin a discussion of existing views from Oza’s (2022)
paper on theories of nonsense. Oza correctly recognises the role that
pretence can play in understanding nonsense. He appreciates the link between Diamond’s discussion of imagination and pretence and notes the importance of these problems for the interpretation of Wittgenstein’s Tractatus. As usual, Oza claims that pretence need not involve an intention to pretend. I think, however, that his position is ultimately untenable.

Oza (2022) maintains that a theory of nonsense must satisfy two requirements. The first one is the engagement constraint, which says that a theory of nonsense should explain how one can use nonsense in certain kinds of reasoning. Oza claims that the austere conception of nonsense fails to satisfy the engagement constraint, since it can only read nonsense ascriptions in a metalinguistic way. The metalinguistic reading says that (7) must be read as (8).

(7) It is nonsense to say that a picture is a fact.

(8) The sentence ‘a picture is a fact’ does not express a thought.

The second requirement is the austerity constraint, which states that nonsensical sentences must not express propositions (or truth-conditions, or thoughts, etc). This constraint is inconsistent with the so-called no nonsense view (Bradley 1978, Prior 1954, Magidor 2009, 2013). According to this account of nonsense, sentences describing category mistakes such as (1) are not nonsensical but are necessarily false. Consequently, (1) expresses the proposition that the pot is drinking coffee and describes the state of affairs that the pot is drinking coffee. However, this proposition is false in every possible world, and there can be no state of affairs that it describes. Oza claims that this view of nonsense implies that the concept of nonsense loses its critical force. If nonsensical sentences express propositions, then they can play the same role as sense. Thus, there is no conceptual difference between nonsense and sense.

---

30 The austere conception of nonsense holds that there is only one kind of nonsense, and that this kind of nonsense has its source in the lack of meaning in one of the components of a sentence. In contrast to the standard view of nonsense, it says that the nonsensicality of (2) is not the result of a violation of the rules of logical syntax. For a discussion, see (Conant 2001, 2002, Dain 2006, 2008, Diamond 1978, 1981, 2005). For a recent defence, see my (Bogucki 2023).
Oza claims that the only view which satisfies both the austerity constraint and the engagement constraint is the pretence account of nonsense which he proposes. His view has it that nonsensical sentences do not express propositions, but that a typical user of language who engages with nonsense pretends that a nonsensical sentence express a proposition. This view satisfies both the engagement constraint and the austerity constraint. A pretence account of nonsense satisfies the first constraint, since an agent operates with a make-believe proposition in reasoning. The second constraint is satisfied because Oza accepts the view that category mistakes are nonsensical. The two constraints are clearly satisfied, so what’s wrong with Oza’s framework?

The main problem with Oza’s position is that it mischaracterises the scope of the theory of understanding nonsense. The austere conception of nonsense, the standard view of nonsense, and the no nonsense view are substantive theories that describe the properties of a class of sentences. These theories debate the status of category mistakes, what semantic properties these sentences have, and what the source of these semantic properties is. The austere conception and the standard view agree that category mistakes are nonsensical, but they differ on how to explain this fact. The no nonsense view claims that category mistakes are meaningful and tries to explain their special semantic status in a different way. A theory of understanding nonsense is a higher-level theory, and it can be reconciled with any theory of nonsense.31 The theory of understanding nonsense aims to explain

31 Any theory that recognises some sentences as nonsensical. The no nonsense view says that category errors are meaningful, but it does not exclude other kinds of nonsense. Jabberwocky cases (nonsense made up of neologisms), terms that have no definition or clear theoretical role (cases discussed by Cappelen), and illusions of thought in general (discussed by Cappelen (2013) and J. A. Keller and L. J. Keller (2021)) are compatible with this view of nonsense. In fact, Cappelen (2013. 86, fn. 5) is quite sympathetic to the no nonsense approach, and he doesn’t see any tension with his other examples of nonsensical sentences. Furthermore, there are simple cases of nonsense (e.g., “Mark is xywwww”, “The cat on the mat is phlump”) that contain a word without meaning. These sentences can be used because they can be the subject of nonsense attributions, e.g., “David believes that the cat on the mat is phlump”. More generally, simple cases of nonsense can also be used as assertions and lead to assertoric nonsense.
how we understand nonsensical sentences, how we can reason with nonsensical sentences, and how to explain the difference between understanding nonsense and understanding sense. As we have seen above, theories of nonsense discuss a different class of problems. One can be a proponent of the austere conception and accept the pretence theory of understanding nonsense. One can embrace the standard view of nonsense and reject the pretence theory of understanding nonsense at the same time. Furthermore, if one accepts the austere view and the pretence theory of understanding nonsense, then both of Oza’s constraints are satisfied. The explanation for this fact is simple: according to the austere view, nonsensical sentences do not express propositions, hence the austerity constraint is satisfied; on the other hand, the pretence theory of understanding nonsense provides an explanation of how engagement with nonsense is possible. Oza’s mischaracterisation of theories of understanding nonsense has implications for his whole framework. In fact, he believes that his view is a ‘pretence account of nonsense’ rather than a theory of understanding nonsense. Next, I will show that Oza’s position also mischaracterises the concept of nonsense itself.

Oza thinks that the pretence account of nonsense is on the same level as the austere view of nonsense, hence he wants to show that the austere view doesn’t satisfy the engagement constraint. His main worry is that, according to the austere conception, ‘when we produce nonsense we aren’t using our conceptual capacities. But our engagement with nonsense draws precisely on these capacities’ (Oza 2022, 7). This has some important consequences for our understanding of language. Firstly, there is no explanation of the difference between sentences (7) and (9).

(7) It is nonsense to say that a picture is fact.

(9) It is nonsense to say that das Bild eine Tatsache ist.

A proponent of the austere view can’t say that the difference is simply knowledge of a language since she must accept the metalinguistic reading of (7) and (9). The metalinguistic reading rejects the appeal to concepts. Secondly, there is no explanation of how a language user understands a sentence that she has never heard before (Oza 2022, 9). The main motivation for the austere view is a reading of Frege’s context principle, which says that the meaning of a word is its contribution to the meaning of the
particular sentence in which it occurs. This reading implies that a word makes no general contribution to a sentence.

First of all, I think the difference between (7) and (9) is understandable if we combine the austere view with the pretence theory of understanding nonsense, and the explanation of this difference is the same as Oza has given. However, Oza’s concern that the austere view is incapable of accounting for our conceptual capacities has some deeper motivations that are also present in the worry about the productivity of our language. These concerns are based on the view that the austere conception is incompatible with the existence of general rules of language. \(^{32}\) I think these worries are based on a misunderstanding of the austere view of nonsense. Historically speaking, Diamond has acknowledged that there are semantic and syntactic rules that tell us how to understand the meanings of words and sentences (Diamond 1978, 199–201, 1981, 19–20). These general rules underlie our understanding of sentences. For instance, if I know the rules of English, then I know that the meaning of the sentence ‘Venus is more massive than Mercury’ is determined by the meanings of ‘Venus’, ‘Mercury’ and the relational expression ‘\(x\) is more massive than \(y\)’ (Diamond 1981, 19). Furthermore, the knowledge of these general rules tells us that ‘Venus is more massive than Mercury’ has some elements common with ‘Venus is less massive than Saturn’, but not with ‘Die Venus ist weniger massiv als der Saturn’. I think that the existence of general rules of language is sufficient to explain the difference between (7) and (9) without resort to a pretence theory of understanding nonsense. (7) is more informative than (9) because it makes use of and requires our knowledge of the general rules of different languages.

Another difference between Diamond (the proponent of the austere view) and Oza (the advocate of the standard view of nonsense) is that Diamond (1981, 19–20) claims that understanding the rules of language is conditional. The conditional understanding of rules means that we can use the general rules of our language to characterise the meaning and structure

---

\(^{32}\) Glock (2004, 2015) and Liptow (2018) have also raised this objection. I cannot discuss the details here, but I have elaborated on this problem and responded to it in (Bogucki 2023, 19–21). The objection has something to do with the Glock’s and Liptow’s reading of the principle of compositionality.
of a sentence, but only conditionally. If a sentence is nonsensical, then there is no way of identifying its syntactic and semantic elements.\textsuperscript{33} I have argued elsewhere that this aspect of the austere view is not incompatible with the principle of compositionality and productivity of language, hence I will limit myself to one observation.\textsuperscript{34} The standard view of nonsense is also committed in some cases to the conditional understanding of rules of language, and the defenders of this view see no problem in appealing to it in the context of the principle of compositionality.\textsuperscript{35} So, the conditional understanding of rules shouldn’t be seen as a problem \textit{per se}. The existence of semantic and syntactic rules underlies our knowledge of language, but sometimes the knowledge of rules tells us that a certain sequence of signs hasn’t been used and should remain without semantic interpretation. Lack of interpretation is not a sign of absence of conceptual abilities.

I have already touched upon another problem with Oza’s framework, namely the problem with ascription of logical form to sentences. In my view, only interpreted signs have logical structure, since meaningful sentences have logical form by virtue of their contents. Nonsensical sentences have no content since they don’t express propositions (or truth-conditions, etc.); therefore, the only principled way of ascribing logical structure to nonsensical sentences is in terms of their surface grammar. The opposite view is a necessary ingredient of the standard view of nonsense, and thus of Oza’s position. Proponents of the standard view ascribe logical form to nonsensical sentences and their constituents, and they maintain that a sentence is nonsensical because of the incompatibility of the logical categories of its

\textsuperscript{33} For example, if we take the meaningful sentence “Venus is more massive than Mercury”, we can determine that, according to the general rules of language, “Venus” refers to the planet Venus, “Mercury” refers to another planet Mercury, and “is more massive” expresses a relation between two objects. However, if the sentence in question is nonsensical, it is not possible to identify the logical elements because the term “Venus” can be used as a predicate in one sentence and as a proper name in another (see Diamond 1979, 209–211, 1981, 7–13, 19–21). The conditional understanding of semantic and syntactic rules fits well with the claim that nonsensical sentences don’t have a logical form.

\textsuperscript{34} See again my Bogucki (2023, 17–23).

constituents. (2) is nonsensical because the logical category of ‘Julius Caesar’ and the logical category of ‘a prime number’ are incompatible. Predicating an abstract property of a concrete object violates the rules of logical syntax. Nonsensical sentences have logical structures; thus, their constituents possess logical categories because the logical properties of sentences are the source of the exclusion of nonsensical sentences from language. As I said in the previous section, an attribution of logical form is problematic because sentences can express different propositions, and these propositions can have incompatible logical forms. But there is another problem with this view that is again common to Oza’s framework and to any standard view of nonsense: it is the problem of making sense that is too similar to nonsense. The challenge is that an explanation of why a sentence is nonsensical leads to attributions of properties to sentences that can only be made about meaningful thoughts and propositions. Oza (2022, 10) claims that the no nonsense view cannot deal with this problem and notes that, according to this view, ‘when a subject attempts to understand a nonsense sentence, they exercise the same capacities they would exercise in understanding the sentence’s words and structure elsewhere’. However, Oza’s position seems to imply the same difficulties. Typically, meaningful sentences express propositions because they consist of concepts and logical forms. According to Oza, nonsensical sentences also consist of concepts and logical forms. His view implies that meaningful and nonsensical sentences share common concepts and logical properties. This similarity can lead one to the conclusion that

I also observed in the previous section that we ascribe logical forms to nonsensical sentences which, in fact, do not exist in our language and are said to be incompatible with logical syntax. If these logical forms are incorrect, then they must also be excluded from sentences.

Oza (2022, 10) formulates this difficulty as follows: ‘(...) there is some difficulty in explaining why a “thought” is meaningless without presupposing that it is meaningful’.

The no nonsense view and Sorensen’s account (2002) blatantly violate the austerity constraint because they claim that nonsense has contents. The standard view of nonsense denies that it expresses propositions, but it does not explain why this is so. It seems as if nonsensical sentences express some illogical thoughts. Hence, there is a reason to think that the standard view non-blatantly violates the austerity constraint.
‘It’s unclear, then, why a subject attempting to understand nonsense fails to grasp a thought: the subject’s activity looks the same as in the good case’ (Oza 2022, 10). James Conant (2001) has argued that such a problem is common to any rendering of the standard view of nonsense, and Oza’s framework inherits this difficulty.

5. The language of thought hypothesis and nonsense

L. J. Keller and J. A. Keller (2021) propose another solution to the riddle of understanding nonsense. In the Introduction, I formulated the riddle in terms of propositions, truth-conditions, and other possible contents of sentences. So, if a sentence doesn’t express a proposition, then it is hard to explain how an agent can understand this sentence. Typically, understanding a sentence consists in understanding the proposition that it expresses. L. J. Keller and J. A. Keller (2021) agree that there are some cases in which an agent believes nonsensical sentences. Furthermore, they argue that such an agent is usually justified in thinking that a nonsensical sentence contains some content. L. J. Keller and J. A. Keller’s explanation of the phenomenon of understanding nonsense (and illusions of thought) is to reject the view that to understand a sentence is to understand the proposition that it expresses. Instead, they accept the language of thought hypothesis.

The LOT hypothesis is said to be a language-like system of internal representation.39 According to it, beliefs and other propositional attitudes are not relations to propositions. All propositional attitudes are mediated. When an agent understands a sentence, she entertains a mental representation of the proposition expressed by this sentence. Propositions and other possible contents of sentences are mediated by our supposed internal language; they don’t appear directly in a language user’s mind since only their mentalese translations do so.40 On the other hand, the contents of sentences

---

39 Fodor (1975) and Schiffer (1981) are the locus classicus of LOT.
40 The mind is divided into a number of boxes because we can have different propositional attitudes. To have a propositional attitude is to have a mental translation of the content of a sentence that describes our attitude (Schiffer 1981).
translated into mentalese determine the contents of our mental actions. L. J. Keller and J. A. Keller (2021) appeal to this feature of LOT in their solution to the riddle of understanding nonsense. They distinguish between ‘beliefs’, ‘thoughts’, etc., understood as some acts of believing, thinking, and ‘beliefs’, ‘thoughts’, understood as some contents of these acts.\textsuperscript{41} When I think the sentence ‘Joe Biden is the current president of the United States’, then there is the proposition that describes the content of my belief, and there is the act of thinking this content. So, what is going on when we are thinking nonsensical sentences? L. J. Keller and J. A. Keller (2021, 242) say:

> With this distinction [between the acts of thinking and their contents] in hand, we can say that while illusions of thought do not, by definition, involve thought contents, they may still involve beliefs, hopes, and desires. It’s just that those beliefs, hopes, and desires—those thoughts—are empty. We can thus give a unified account of thinking, speaking, and writing nonsense: thinking nonsense involves actually thinking empty thoughts; speaking nonsense involves actually speaking empty words, and writing nonsense involves actually writing sentences that don’t (actually) express contents.

L. J. Keller and J. A. Keller (2021) maintain that mentalese representations of nonsensical sentences have some contents, but these contents are empty. Every public language sentence has its mentalese counterpart, and the proposition expressed by the public language sentence determines the content of its mentalese counterpart. Nonsensical sentences are devoid of content, hence when they are represented in mentalese sentences, they don’t give them any specific content. On the other hand, the act of thinking about a content is different from the content itself, so there is an act of thinking about a nonsensical sentence even though that sentence has no content. Does this solve the puzzle of understanding nonsense?

First of all, if someone doesn’t accept the language of thought hypothesis for some other reason, the price of solving the puzzle of understanding is

\textsuperscript{41} L. J. Keller and J. A. Keller claim that the act of believing or thinking is a concrete mental state or event. They mark these states with the subscript \textit{s}.

\textit{Organon} F 30 (4) 2023: 372–411
quite high. LOT is a controversial view that proposes a substantive theory of mind and a non-classical account of propositional attitudes. It carries with it a number of ontological commitments. One may raise the concern whether it is worth appealing to LOT in order to solve a particular problem in the theory of language. Moreover, even if LOT is theoretically indispensable, it needs empirical confirmation, and the tests may turn out to be unsuccessful. Of course, LOT also has some important advantages, but it seems to me that the pretence theory of understanding nonsense has fewer and less serious commitments. For this reason, it is preferable to LOT, even if both views are equally successful in explaining the problem of understanding nonsense.

So, is LOT able to explain the riddle of understanding nonsense? In my opinion, the answer to this question is negative. Firstly, it is hard to understand what translations of nonsensical sentences into mentalese are. L. J. Keller and J. A. Keller (2021) claim that if a sentence is nonsensical, then its mentalese translation is also without meaning. (3) is nonsensical, so one’s belief box contains its mentalese counterpart ‘THE THEORY OF EVOLUTION IS EATING BREAKFAST’. The explanans, namely the fact that one’s belief box contains a nonsensical mentalese sentence, doesn’t seem any more intelligible and self-explanatory than the explanandum – the fact that one believes a nonsensical sentence. What are mentalese translations of nonsensical sentences? L. J. Keller and J. A. Keller (2021) maintain that mentalese translations of nonsenses have contents that are empty. This explanation doesn’t tell us much. Do mentalese translations also have empty logical contents? The notion of empty contents seems no more informative than the well-known concept of sentences without contents. It is hard to see any theoretical gain in postulating nonsensical mentalese sentences. Therefore, it seems to me that L. J. Keller and J. A. Keller (2021) should have spent more time elaborating on the utility of this concept.

Secondly, one might doubt whether there really are such things as mentalese translations of nonsensical sentences. According to LOT, to believe that ‘snow is white’ is to have a mentalese translation of the content of the sentence ‘snow is white’ in one’s belief box. So, if I believe that \( p \), I have a mental

---

42 It’s L. J. Keller and J. A. Keller’s (2021) convention to mark mentalese counterparts with capital letters.
representation of a proposition that $p$. The problem is that nonsensical sentences don’t express propositions (or truth-conditions). Hence, there is no content to be translated into mentalese at all. What do we translate into mentalese in such a case? In the previous paragraph, I assumed that we could ignore this problem and postulate that we translate sentences, but sentences don’t have semantics. They have to be interpreted, and they can only get content thanks to propositions or truth-conditions. LOT’s mentalese sentences are supposed to have semantics, but they can’t get them from semantically uninterpreted sentences. I think that some proponents of LOT are aware of this problem (Recanati 1997, Sperber 1997). Dan Sperber (1997) claims that one’s belief box can only contain a meta-representation of the proposition $p$ if one doesn’t fully understand $p$. So, if a schoolboy doesn’t understand the proposition that Cicero’s prose is full of ‘synecdoches’ (because he doesn’t understand the meaning of the term ‘synecdoche’), his belief box can contain a meta-representation of this proposition but not its representation.43 Recanati (1997, 91) agrees that ‘(...) a sentence cannot make its way into the mind (whether into the belief box or elsewhere) if it contains uninterpreted symbols’. A representation must be interpreted in some or other way, otherwise it is difficult to understand what it means to mentally entertain uninterpreted symbols. Half-understood sentences (‘Cicero’s prose is full of “synecdoches”’) and nonsensical sentences are uninterpreted.

Recanati appeals to Kaplan’s character-content distinction to solve the problem of understanding uninterpreted symbols44. Sometimes a representation of a proposition is determined by the character of an expression. Character is informative independently of content since it shows the route to the proposition, even if that proposition doesn’t exist. Recanati postulates introducing a deferential operator which provides character to uninterpreted sentences. The deferential operator $R_x()$ applies to the symbol $\sigma$ and yields a complex expression $R_x(\sigma)$. The character of $R_x(\sigma)$ takes us from a context in which the speaker tacitly refers to a particular cognitive agent

43 The schoolboy believes that the teacher says that Cicero’s prose is full of ‘synecdoches’, but he can’t entertain the proposition that Cicero’s prose is full of ‘synecdoches’.

44 See Kaplan (1989) on the distinction between character and content.
to the content which $\sigma$ has for $x$. This solves the problem, since one’s belief box doesn’t have to contain the representation of the uninterpreted sentence or its meta-representation. This belief box contains a deferential representation that is the proposition-like character of the previously uninterpreted sentence. For instance, the character of the sentence ‘Cicero’s prose is full of “synecdoches”’ takes us from the cognitive agent that is tacitly referred to (i.e., the teacher) to the content that the symbol ‘synecdoche’ – and thus the whole sentence – has for the teacher. Loosely speaking, the character says that the content of ‘synecdoche’ is whatever concept the teacher was referring to, and that concept – together with the rest of the constituents – determines the sentence.

In my opinion, Recanati’s amendment is only plausible for sentences which have some natural candidates for a particular character of an expression. Typical cases of nonsense, like ‘Julius Caesar is a prime number’, don’t easily submit to the character/content distinction. Another of Recanati’s examples is the sentence ‘The unconscious is structured like language’, the character of which again points us to the author’s intended concept. The difference is that, according to Recanati, the author of this *bon mot*, Jacques Lacan, didn’t have a specific meaning in mind when he made this statement. Hence, the sentence has a character, but it doesn’t have any content. The character of the sentence allows us to get the deferential representation. It is not clear what the character of the sentences ‘The pot is drinking coffee’ or ‘Julius Caesar is a prime number’ is supposed to be. In both cases, the utterer of the sentence wanted to express a proposition that doesn’t exist. Character can’t direct us to the proposition because the utterer knew that the proposition is semantically defective. The riddle of nonsense is to find the proposition that is missing. Recanati’s theory doesn’t say how to find the lacking proposition, nor does it provide a plausible character for the most important cases of nonsense. His two examples are similar in that they emphasise the author’s intention, but there is no such authorial intention in the most interesting cases. Furthermore, the community of users does not determine the expressed proposition. There is no publicly available

---

45 Strictly speaking, the sentence is semantically defective and fails to express a proposition.
meaning. The character of nonsense can’t refer to a non-individualistically interpreted user. Thus, Recanati’s solution may seem promising for half-understood sentences, but it fails in the case of full-fledged nonsense. Full-fledged nonsense doesn’t have plausible characters, and it’s even hard to conceptualise some loose descriptions of the candidates.

Finally, L. J. Keller and J. A. Keller’s (2021) account cannot explain the role of nonsensical sentences in reasoning. Mentalese translations of nonsensical sentences inherit semantic properties from their contents. On the one hand, nonsensical sentences have no content, and mental translations can’t get semantic properties from them. On the other hand, nonsensical sentences don’t have logical relations and logical structures. Their mentalese translations should have the same properties, which is why it is hard to understand how we can use nonsensical mentalese sentences in reasoning and draw conclusions from them. Does the mentalese sentence ‘the pot is drinking coffee at t₁’ entail the mentalese sentence ‘the pot is not thirsty at t₂’? As we have seen, in some circumstances nonsensical sentences appear in inferential connections and justify our actions. Moreover, L. J. Keller and J. A. Keller discuss sentences like ‘Witches cast spells’ which are also semantically defective. People think that if witches cast spells, then witches are dangerous, and witches should be avoided since they can cause numerous misfortunes. The properties of empty mentalese sentences don’t explain why these sentences are logically related. The explanation provided by nonsensical mentalese sentences does not seem to be any more helpful in such cases than their linguistic counterparts. The distinction between the act and the content of thinking doesn’t help either. The act of thinking nonsensical sentences has empty content that can play no role in reasoning. The fact that mental realisations of this content are distinct doesn’t change much.

---

46 Contrast this with the half-understood sentence ‘I have “arthritis” in my thigh’, discussed by Recanati. In this case, a speaker doesn’t know the public meaning of this sentence, but the language community collectively determines the content (Burge 1979).

47 See L. J. Keller and J. A. Keller (2021, 227, fn. 6) for an explanation of the semantic defectiveness of this sentence.
6. Conclusion

In this paper I have explored the riddle of understanding nonsense. This riddle concerns the possibility of entertaining nonsensical sentences, judging their content, and drawing conclusions from them, given the fact that nonsensical sentences don’t express propositions that are standard objects of propositional attitudes. In Section 2, I elaborated on the importance of this riddle. In Section 3, I proposed a pretense theory of understanding nonsense. In Section 4, I examined Oza’s alternative framework, which is based on the standard view of nonsense and attributes logical forms to sentences. In Section 5, I discussed L. J. Keller and J. A. Keller’s proposal to explain the riddle of understanding nonsense by means of the language of thought hypothesis. I think my considerations show that Oza’s and L. J. Keller and J. A. Keller’s proposals have some shortcomings that make my own view more favourable than their positions.

Disclosure statement

No potential conflict of interest was reported by the author.

Acknowledgements

I would like to thank Tadeusz Ciecierski, two anonymous referees for this journal and the audiences at the MLAG Seminar at the University of Porto, the Seminar Series in Analytic Philosophy at the University of Lisbon and the 11th Polish Philosophical Congress at the University of Łódź for their helpful feedback.

Funding

This work was supported by the National Science Centre, Poland under Grant 2022/44/C/HS1/00054.

References

https://doi.org/10.1007/s11098-012-9915-6


Organon F 30 (4) 2023: 372–411


