

Philosophia  
<https://doi.org/10.1007/s11406-020-00198-2>



# Ineffability: the Very Concept

Sebastian Gäb<sup>1</sup> 

Received: 10 October 2019 / Revised: 21 January 2020 / Accepted: 21 February 2020

Published online: 14 March 2020

© The Author(s) 2020

## Abstract

In this paper, I analyze the concept of ineffability: what does it mean to say that something cannot be said? I begin by distinguishing ineffability from paradox: if something cannot be said truly or without contradiction, this is not an instance of ineffability. Next, I distinguish two different meanings of ‘saying something’ which result from a fundamental ambiguity in the term ‘language’, viz. language as a system of symbols and language as a medium of communication. Accordingly, ‘ineffability’ is ambiguous, too, and we should make a distinction between weak and strong ineffability. Weak ineffability is rooted in the deficiencies of a particular language while strong ineffability stems from the structure of a particular cognitive system and its capacities for conceptual mental representation. Mental contents are only sayable if we are able to conceptualize them and then create signs to represent them in communication.

**Keywords** Ineffability · Language · Communication · Signs

Ineffability is a peculiar subject. It comes up occasionally in a variety of areas like philosophy of language, logic, philosophy of religion, or aesthetics, but until recently, it has rarely been itself the focus of attention.<sup>1</sup> We stumble upon the idea in Laozi and Plotinus, Adorno and Frege, but it is hard to see the one thread which holds all these works and debates together. My intention in this paper is to find this thread by unraveling the very concept of ineffability. I take it that when we say that something is ineffable, we mean that it cannot be said – whatever this means. So, if we want to understand what it means that something cannot be said, we will have to understand what it is we are unable to do. Therefore, my starting point will be the concept of saying something. I will show that this concept is fundamentally ambiguous and that we

---

<sup>1</sup>Notable exceptions are the book-length studies of Jonas (2016), Bennett-Hunter (2014), and Kukla (2005); also A.W. Moore’s work on the subject scattered over roughly a dozen or so papers and book chapters.

✉ Sebastian Gäb  
s.gaeb@lmu.de

<sup>1</sup> Faculty of Philosophy, Philosophy of Science and Religious Studies, LMU Munich, 80539 Munich, Germany

should accordingly distinguish two different yet related meanings of ‘ineffable’. In what follows, I will confine myself to analyzing the concept of ineffability, i.e. I will not touch upon the question whether or not there actually are ineffable things and what evidence we might have for their existence. Any discussion on the possibility of ineffable objects requires a preliminary investigation of the concept itself – so that we understand what it is the existence of which we want to affirm or deny.

## 1 What Ineffability is Not

If something is ineffable, it cannot be said. This much is certain. But before we start analyzing what this means, let me briefly point out what it does not mean: being ineffable is not the same as being false, not even necessarily false or paradoxical. Just because something cannot be said *truly* does not mean that it is ineffable. Let’s take a look at the following example of this subtle confusion (found in Shaw 2013) to explain this difference. Shaw states that a proposition is ineffable if ‘there is no expression of L which expresses that [proposition].’ (Shaw 2013, 65) One of his examples is:

- (1) ‘Let  $n$  be the integer one greater than the greatest integer I successfully define today.’ (Shaw 2013, 73)

This sentence seems to imply a perfectly meaningful proposition:

- (2) There is an integer  $n$  one greater than the greatest integer I successfully define today.

But upon reflection, we realize that (2) is not perfectly meaningful at all: Whatever other integers I may have defined today, utterance (1) also contains a definition of an integer and if it defines any integer at all, it will be an integer one greater than itself, which is clearly impossible. So, the proposition ‘ $n$  is the integer one greater than the greatest integer I successfully define today’ cannot be expressed, since as soon as I utter a sentence expressing it, it will be false.<sup>2</sup> Note that this depends on me making the statement in question – if someone else says ‘ $n$  is the integer one greater than the greatest integer *you* successfully define today’, no paradox arises. This might be a hint that what Shaw describes is not a genuine case of ineffability. After all, the proposition is not ineffable per se (it’s right there, on page 79). It is just impossible to express this proposition without contradiction. It can be said, it just cannot be *asserted*. Shaw’s examples resemble instances of performative contradictions like ‘I am sleeping’. The proposition that you are currently sleeping is perfectly fine and can be stated without problems, it is just impossible to utter it and say something true. Likewise, you cannot utter proposition (2) without thereby rendering it false. Shaw (2013, 85) objects that trying to express an ineffable proposition (in his sense) does not alter the truth value of the proposition. Rather, the proposition cannot be expressed *simpliciter*. But what does this mean? Take the following proposition: there is a number  $n$  which is one greater than the greatest number which has ever appeared in a philosophical paper. Pending further research on the matter,

<sup>2</sup> Kukla (2005), 149ff. describes the similar case of ‘unreportability’. My objections apply to him as well.

let us assume that this number is 1,000,000. Then  $n$  will be 1,000,001. But now that 1,000,001 has appeared in a philosophical paper,  $n$  has to be 1,000,002, and since this number has appeared,  $n$  has to be 1,000,003 and so on. I will never be able to state  $n$  correctly in this paper. But does this mean that the proposition that there is a number  $n$  which is one greater than the greatest number which has ever appeared in a philosophical paper cannot be expressed? No. It is expressed by the very words you just read. You could even say which number it is: assuming it really is the number one greater than 1,000,003, just open your mouth and say the words ‘The greatest number which has ever appeared in a philosophical paper is...’ replacing ‘...’ with the number one higher than 1,000,003. As long as it does not appear in a philosophical paper, this proposition is perfectly expressible and true. The problem is just that once the number has been mentioned in a philosophical paper, nothing can ever make this proposition true. Uttering the proposition (in the right context) has an effect on what could possibly count as truth-maker for the proposition – nothing, since the utterance generates an infinite regress of spurious truth-makers. And if nothing can make a proposition true, it will necessarily (given a two-valued logic) be false. So, uttering the proposition does not make it false (unlike ‘I am sleeping’), but uttering the proposition destroys any chance the proposition might ever have of being true. But this still does not mean that proposition (2) is ineffable. It just means that you cannot express proposition (2) and be right about it. Admittedly, it is not unusual in philosophical literature to understand ‘ineffable’ as ‘impossible to say truly (or without contradiction)’,<sup>3</sup> and if this is common usage, we must simply accept that (who are we to tell others how to speak?). But we should nevertheless be aware that this meaning of ‘ineffable’ differs from other cases like mystical or aesthetical ineffability and bears much more resemblance to liar-like paradoxes. So, for reasons of clarity, I suggest to separate ineffability from any questions about truth and falsehood: saying that a proposition is ineffable implies nothing about the truth or falsity of this proposition.

## 2 Saying Something

What does it mean when I tell you that something cannot be said? Obviously, this depends on what it means to say something. Well, saying something is what language does or what we do with language, and so asking for the meaning of saying something is in a way tantamount to asking: what is language?

This question is far too general for a simple answer, but even if we ignore the cultural, social or literary dimensions of language and focus on the bare concept of language itself, a fundamental ambiguity remains. The simplest definition of language is probably something like this: a system of symbols that is used to communicate.<sup>4</sup> So, language clearly has two essential properties: it is a system of symbols governed by rules pairing those symbols with meanings. And it is what we use to convey thoughts. Accordingly, language can be regarded in two different ways: (a) as a system of symbols which can be combined following certain rules to produce other symbols and which are assigned certain values (called meanings, or, more technically, *semantic values*) following other rules; or (b) as a medium of communication, i.e. a certain type

<sup>3</sup> Besides Shaw and Kukla, see e.g. Alston (1956), 507 or Kellenberger (1979), 310ff.

<sup>4</sup> See, for example, Lyons (1981), 8.

of observable human behavior which is used to express thoughts or cause actions.<sup>5</sup> And depending on which aspect we prefer, our answer to the question ‘What is language?’ (and ‘What is ineffability?’) will be a different one.

If we focus on the first aspect, we will emphasize the systematic and logical character of language (call this the *logical* model of language). According to this model, there are atomic and molecular units of meaning, called terms or expressions, which can be combined to produce larger units, called sentences or propositions. The rules that guide this process of combination are called the syntax of the language and the meaning of these larger units results from its semantic rules. The proposition’s individual elements have certain semantic values and refer to certain parts of reality while the proposition itself refers to a state of affairs. Thus, the proposition as a whole has a truth-condition which is its meaning: ‘To understand a proposition is to know what is the case if it is true.’ (Wittgenstein 1922, *Tractatus* 4.024) According to this paradigm, ‘saying something’ means being an element (proposition or term) of a certain language.  $a$  or  $F$  says something if and only if there is a syntactically correct expression  $F(a)$  in some language  $L$  which has some meaning given the semantic rules of  $L$ . According to this paradigm, it is the expression itself which says something.

If on the other hand we emphasize the second aspect, we will see language as an instrument of communication and a specific type of human behavior (call this the *communication* model). On this model, linguistic utterances are always embedded in contexts of time, place, and action and the meaning of an utterance (or its *assertoric content*<sup>6</sup>) is derived from the specific use of the symbols I employ and their respective semantic values in the context in question. ‘The meaning of a word is its use in the language.’ (Wittgenstein 2009, *Philosophical Investigations* § 43) Understood this way, ‘saying something’ means doing something. It’s not that the expression says something, but it’s that I say something via the expression. The meaning of an expression is determined by the speaker’s intentions or the communicative function it fulfills in a given situation.

<sup>5</sup> As David Lewis puts it: ‘I distinguish two topics: first, the description of possible languages or grammars as abstract semantic systems whereby symbols are associated with aspects of the world; and second, the description of the psychological and sociological facts whereby a particular one of these abstract semantic systems is the one used by a person or population. Only confusion comes of mixing these two topics.’ (Lewis 1970, 19)

<sup>6</sup> I am grateful to an anonymous reviewer for drawing my attention to the distinction between semantic value and assertoric content, which roughly coincides with my distinction between the logical and the communication model (see Ninan 2010; Rabern 2012 for details). However, even if it may be helpful to explain the basic idea in these terms, the two distinctions are not identical. The distinction between semantic value and assertoric content, too, reflects the different meanings of ‘language’ (what our sentences say and what we mean by them), but on a different level. Their relation (between semantic value and assertoric content) is usually understood like this: the semantic values of each expression and the recursive rules of combination together yield a sentence’s semantic value, which in turn determines the assertoric content it has (or would have) in a given context (Rabern 2012). In this picture, we derive assertoric content from a sentence’s semantics and the state of the world (and therefore sometimes speak of ‘postsemantics’). On my account, though, what is said can only be determined by semantics (plus postsemantics) alone if we work on the level of a theory of meaning, not of meaningfulness (see the following paragraph). Or, to put it differently: you don’t need intentionality for assertoric contents. So, the concept of assertoric content, as usually understood, is too narrow. But if we understand assertoric content in a very wide sense in which it is not just determined by semantic values, but by a speaker’s intentional mental states which *inform* these semantic values, we may put the distinction between the two models of language in these terms.

If we transpose this distinction from language to meaning, it corresponds to a distinction between theories of meaning and theories of meaningfulness.<sup>7</sup> A theory of meaning states the general rules which determine the meaning of some symbol or sign (e.g.: the meaning of a sentence is its truth-condition) and explains how to pair expressions and their semantic values. For these theories, the fact that the symbols in question have meaning is given. All the theory needs to explain is what the meaning of a specific sign is by stating the principles which determine the attribution of meaning to symbols. It does not touch on the deeper question of how these symbols come to be carriers of meaning in the first place. This is the task of theories of meaningfulness: explaining how some arbitrary (and in itself meaningless) physical object or event becomes a sign, a carrier of meaning. While a theory of meaning states that the meaning of a sentence is its truth-condition, the theory of meaningfulness explains why sentences and truth-conditions are tied to each other at all (for why does a concatenation of characters like ‘it’s raining’ have a truth-condition, but ‘agg beware’ doesn’t?).<sup>8</sup> Both theories supplement each other (Lewis 1975, 35), since both aspects are mutually dependent, but the concept of language is still ambiguous. Saying something is either a property of certain elements of a system of symbols or a type of action performed by persons.

### 3 Two Meanings of *Ineffable*

This, of course, has consequences for the concept of ineffability: we should expect ineffability to exhibit the same distinction we encountered in the concept of language. Therefore, we need to distinguish two different meanings of ‘ineffable’: (a) weak ineffability or inexpressibility. What is weakly ineffable is ineffable in terms of a theory of meaning: a given language L does not contain any expression that (according to the semantic rules of L) designates what is to be expressed. (b) Strong ineffability or incommunicability. If something is strongly ineffable, it is ineffable in terms of a theory of meaningfulness: it is impossible to create a sign which carries the meaning required to express the ineffable. As a consequence, what is strongly ineffable cannot be communicated.<sup>9</sup>

<sup>7</sup> The terms are borrowed from Cummins (1989), 10 f. Speaks (2014) makes an essentially identical distinction between ‘semantic theories’ and ‘foundational theories of meaning’.

<sup>8</sup> Incidentally, this is the reason why Wittgenstein’s or Grice’s theories should not be seen as alternatives to a truth-conditional theory of meaning (though I’m not saying that maybe Wittgenstein or Grice did not themselves believe this). Both aim at explaining how linguistic signs come to be carriers of meaning (viz. by being used by conscious subjects with certain intentions), not at stating their meaning. They are theories of meaningfulness, not of meaning.

<sup>9</sup> Hofweber (2005), 172 f. makes a similar distinction between language expressible and speaker expressible: a property P is language expressible in some language if there is a predicate that expresses P in every context, and speaker expressible if there is a predicate and a context such that an utterance in the given context expresses P. Depending on whether the relevant context actually exists or is merely conceivable, Hofweber distinguishes further between factually and loosely speaker expressible. But although this distinction resembles mine, they are not identical. If something is language inexpressible or factually speaker inexpressible, it is weakly ineffable in my terminology. If something is loosely speaker inexpressible (which is not possible according to Hofweber), I would call it strongly ineffable. Cf. also Spackman (2012), 304, who also distinguishes semantic (=weak) and communicative (=strong) ineffability, but doesn’t elaborate further on the distinction.

### 3.1 Weak Ineffability (Inexpressibility)

(I<sub>1</sub>) A proposition  $p$  is weakly ineffable in a language  $L$  if and only if no sentence which expresses  $p$  is an element of  $L$ .

This definition raises two questions: (a) what counts as a language and (b) when does a sentence express a proposition?

- (a) On the logical model, languages are finite or infinite sets of expressions or sentences of finite length.<sup>10</sup> Sentences have molecular structure: they are composed of a set of elementary symbols we call words (in natural languages) or terms. Symbols in natural languages are either temporal series of sounds or spatial series of written characters. Every language has syntax, a system of rules defining correct and incorrect combinations of symbols. It also has semantics, a second system of rules which attributes meaning (semantic values) to the elementary symbols and their combinations. The meaning of the complex symbols (sentences) is a function of the meaning of the elementary symbols which constitute it. Moreover, every sentence<sup>11</sup> (and only sentences) has a truth-value, which is the product of its meaning (as determined by the language's semantics) and the facts obtaining in the world. So, a language can be described as a set of ordered pairs of symbols and meanings (Lewis 1975, 3). What then is the meaning of a given sentence  $\sigma$  of  $L$ ? That which determines its truth-value. The meaning of a sentence is its truth-condition, i.e. a function from a possible world to a truth-value.<sup>12</sup> For example, the meaning of 'Beijing is the capital of China' is a function which renders the truth-value of this sentence for the actual world as true. According to this account, a language  $L$  is a formal system consisting of a (finite) set of symbols and two (finite) sets of rules, syntax, and semantics. This is not so much an exhaustive definition of language but rather the minimal condition of languagehood: nothing could count as language which doesn't exhibit these features.
- (b) What does it mean to express something? It is surprisingly difficult to find explicit answers to this question in the philosophy of language; most philosophers take the notion more or less for granted. A notable exception is A.W. Moore who defines expression as follows:

$x$  expresses  $y$  if and only if (i)  $x$  is a linguistic item with content that makes it either true or false, (ii)  $y$  is a non-linguistic item with content that makes it either true or false, and (iii) the content of  $x$  entails the content of  $y$ . (Moore 2003, 173)

For example, the sentence 'grass is green' ( $x$ ) expresses that grass is green ( $y$ ), because it is a linguistic item with a certain content (that grass is green) which implies the content of, say, the belief that grass is green. (This seems a little odd – doesn't the

<sup>10</sup> I'm following Chomsky (1957), ch. 2.1 and Carnap (1937), 4ff. here.

<sup>11</sup> Or not, depending on whether or not we accept the principle of bivalence for the given language.

<sup>12</sup> If you don't like possible world semantics, nothing in my argument really depends on it.

sentence first and foremost express the *fact* that grass is green rather than my belief that grass is green? And this fact does not have any content, it *is* the content.) The crucial point is (iii): what does it mean that the content of  $x$  entails that of  $y$ ? As Moore explains elsewhere (Moore 1997, 16), entailment is not the same as logical implication. Rather, we should be able to tell a story about  $x$  which has  $y$  as its consequence (Moore keeps this deliberately vague). But which story could that be? It cannot simply be that  $x$  *logically* entails (or implies)  $y$ , since it is impossible to understand implication without understanding expression first: ‘ $x$  implies  $y$ ’ means that it is impossible for  $x$  to be true while  $y$  is false. But to decide whether there could be a case in which  $x$  is true and  $y$  is false, I have to know under which conditions  $x$  and  $y$  are true, which is the same as knowing their meaning. And if I know what  $x$  means, I surely know what  $x$  expresses, too – most notably that it expresses  $y$ . On the other hand, the story cannot be a story of purely causal connections, either. If you tell me that you went to New York last fall ( $x$ ) and I (as a consequence) now want to eat a pastrami sandwich ( $y$ ), then obviously what you told me doesn’t express my desire for pastrami sandwiches. The key to a solution lies in the phrase ‘... that makes it either true or false’: expressions have contents with truth-conditions and there is a systematic relation between the expression and these truth-conditions. Relying on the minimal conditions of languagehood I described above, we could define the concept of expression like this: trivially, a sentence  $\sigma$  of  $L$  expresses a proposition  $p$  if and only if it means that  $p$ . This is equivalent to saying that  $\sigma$  expresses  $p$  if  $\sigma$  is made true (according to the semantic rules of  $L$ ) only by the fact (or state of affairs) that  $p$ .<sup>13</sup> For example, the sentence ‘Lemons are yellow’ means that lemons are yellow by expressing the proposition *lemons are yellow* which is made true only by the fact that lemons are yellow. The proposition that lemons are yellow is therefore expressible in English. So, to express something means to stand in a systematic relation of entailment to something which has truth-conditions. Which proposition is expressed by a sentence depends on its components. In general, any sentence  $F(a)$  will express that the concept  $F$  applies to the object referred to by  $a$ . So, a proposition  $p$  is expressible in  $L$  if and only if a sentence  $s$  can be construed from the symbols given in  $L$  whose truth-condition is derived from the terms that constitute the sentence and which is identical to the truth-condition of  $p$ : it can only be made true by the fact which makes  $p$  true.<sup>14</sup> Note that we do not need to explain expression through meaning (although these two go hand in hand – what is expressed *has* meaning). Expression presupposes only the availability of concepts and a notion of truth:  $x$  expresses  $y$  if  $x$  has a propositional content which is made true by the fact that  $y$ . For every sentence of  $L$ , it is possible to state what it expresses by deriving its truth-condition from its constituent parts. And whether or not some state of affairs  $y$  obtains can be observed without understanding any sentence that expresses this state of affairs. We can understand that some sentence expresses  $y$  without knowing what  $y$  means.

Weak ineffability, then, is a phenomenon on the level of theories of meaning, not meaningfulness. Weak ineffability is a deficiency of a symbolism whose stock of symbols and rules of combination don’t allow us to produce a sentence which expresses

<sup>13</sup> Nothing hinges on the exact ontological status of truth-makers as long as they are distinct from the propositions they make true.

<sup>14</sup> Alternatively, we might put it like this: a proposition  $p$  is expressible in some language  $L$  iff it is possible to form a sentence in  $L$  with the assertoric content that  $p$ .



the ineffable proposition. No sentence can be produced in L (given its symbols and their semantics) which is made true by the inexpressible state of affairs. For example: imagine a rudimentary mathematical language which only contains symbols for natural numbers and the operations of addition and subtraction. In a language like this, the propositions ‘ $5 - 7 = -2$ ’ and ‘ $3 \times 3 = 9$ ’ are inexpressible, since symbols like ‘-2’ and ‘x’ are neither syntactically nor semantically defined for this language.<sup>15</sup> If all we have is the natural numbers, ‘+’ and ‘-’, it is impossible to express the propositions just mentioned. That doesn’t make these propositions ineffable per se – they are easily expressible in a stronger symbolism (like our actual mathematical language). It also doesn’t make them false or meaningless – they are true under precisely defined circumstances and therefore have truth-conditions and meanings. The language just lacks the means to successfully put these propositions into words. Weak ineffability is *language-ineffability*: it results from the shortcomings of a particular language and is therefore in principle removable through another language with greater expressive powers.<sup>16</sup>

### 3.2 Strong Ineffability (Incommunicability)

(I<sub>2</sub>) A proposition p is strongly ineffable if and only if for any subject S: S can be in a mental state that p and S cannot communicate that p.

The crucial question here is: What does it mean to communicate and how does it differ from expressing something?<sup>17</sup> Basically, any act of communication between a speaker S and a receiver R requires that (a) S makes an utterance which conveys information intended by S from S to R and (b) the information conveyed is coded in signs.<sup>18</sup>

Information, as I understand it here, is a purely statistical phenomenon without meaning. Take, as an example, the light on my oven: the light is on as long as the temperature remains below a certain threshold; once the threshold is crossed, it is switched off. This means that in this particular system, two states are possible: either the threshold is crossed or not. The light transmits information about which of these two states obtains. Information, then, is nothing but a decision between two otherwise equally probable states. What these states are or mean is irrelevant. Information is just a measure of binary decisions that must be made to identify a certain state among a set of other possible states. It does not tell us anything about these states unless we interpret it.

<sup>15</sup> For a more elaborate example along these lines see Henle (1949).

<sup>16</sup> It is an open question whether this holds for any sentence. If it does, there will be no absolute ineffabilities and every inexpressible proposition could in principle be expressed in some other language (possibly in one total language). Although I believe this conclusion to be correct, I will not argue for it here as this exceeds the scope of the present paper. But see Hofweber (2005) for an elaborate argument for what he calls the ‘principle of expressibility’.

<sup>17</sup> I will confine myself to linguistic communication, since this is clearly the only type of communication which is relevant for discussing ineffability.

<sup>18</sup> These are only conditions for possible, not for successful communication. Any utterance that fulfills these conditions can be used to try to communicate something, but to make sure that something is actually communicated further conditions have to be satisfied. For example, the receiver has to grasp the information by and because of understanding the speaker’s intention (the conditions detailed in Grice 1957). But since successful communication is merely a special case of communication per se, and since our question is what ineffability is, we only need to focus on the conditions for the possibility of communication. If communication fails because the receiver isn’t paying attention, this is not a case of ineffability.



Thus, if something is incommunicable, it cannot be so due to a failure in transmission of information, since information is just a series of intrinsically meaningless physical events. As long as we are capable of causing events, we are capable of producing information. If something is incommunicable, the reason is that we are unable to interpret the events carrying information as signs.

Transmission of information is a purely physical process. The series of events from the thermostat reaching a certain temperature to switching off the light on the oven can be explained fully in terms of causal links. It doesn't matter what the light being on or off *means* to understand the process. The situation is different if I tell someone to take a seat and the person proceeds to sit down. In this case, it doesn't suffice to list the physical events and their causal connections to explain what happened. It is not the soundwaves I emit that lead to the other person sitting down, but the meaning these soundwaves carry. Had I said '请坐!' instead, nothing (or something else, presumably confused stares) would have happened, even though the exact same soundwaves would have the intended effect on a person who speaks Chinese. The crucial difference between transmission of information and linguistic communication is that communication is not merely physical, but requires the use of signs as a further, non-physical element. Communication is the ability to use linguistic signs, and explaining why we are capable of using signs at all is the task of a theory of meaningfulness. Incommunicability, then, is the inability to signify something: to create a sign for the intended meaning. But how could this be possible? To understand this, we need to take a closer look at the mechanics of signs.

Peirce classically defines signs as follows:

A sign [...] is something which stands to somebody for something in some respect or capacity. (Peirce 1931, 2.228)

Signs are somethings, i.e. they are physical objects, not mere abstractions. But not every physical object is a sign, although any object could be a sign. What makes it a sign is the fact that it 'stands for' or represents something. This is essential and the core element of any theory of signs – nothing could be considered a sign if it does not represent something. But while everyone agrees that signs represent what they are signs for, it is far less clear what it actually means for a sign to stand for something. Most theories agree that there is a two- or three-place relation between the significant (the physical instantiation of the sign), the significate (its meaning) and the referent (the object the sign refers to). (Dyadic or two-place theories drop one of these elements, usually the referent.) Take, for example, the word 'rose'. In this case, the significant will be the concrete physical symbol, the ink on the paper, or the waves of sound emitted; the physical form is arbitrary. The significate is whatever I understand if I understand the sign, or it is that by which the sign as an individual physical entity can stand for something else – in short, its meaning. In our example, the significate would be the abstract concept *rose* or the mental representation ROSE. The process of signification in which a meaningless physical object becomes a meaningful representation of something else is the essential aspect of a sign, and it is also the point of failure in the case of incommunicability. There is no direct connection between the physical sign and the object bypassing the significate: natural signs like smoke for fire or fever for illness are not signs in the strict sense

of the word, because they don't signify anything. The smoke doesn't stand for the fire, it is caused by fire.<sup>19</sup>

A sign represents something by establishing a relation between two objects, the significant and the referent. This relation is mediated by the significate which gives the sign its meaning and in virtue of which the significant means the referent. Any meaningless physical object becomes a sign if we attribute meaning to it. But what does it mean to attribute meaning to an object, for it to be meaningful at all? We don't need to develop a fully-fledged theory of meaningfulness here. To understand the notion of incommunicability, it will suffice to emphasize one essential aspect of any such theory: clearly, the meaningfulness of a sign depends on a subject's intentionality and their mental states. A being is only capable of signification and linguistic representation if it is also capable of mental representation (in the widest sense of the term). Imagine a world without conscious, intentional beings. Even if in such a world the whistling of the wind might sometimes by sheer accident produce the sound 'wine', obviously these sounds would not mean anything. The meaningfulness of signs presupposes intentional states, or, to put it another way: linguistic meaning depends on mental meaning. Signs have meaning and draw their power to represent something from mental representations.<sup>20</sup> This is not meant as a commitment to intention-based semantics, although any intention-based semantics will have to accept this point as given. What Grice and his followers had in mind was to specify how linguistic meaning is reducible to speaker intentions; all I claim here is that linguistic signs are impossible for entities which lack any ability for mental representation, no matter how the specific meaning of these signs is constituted (intentions are a necessary condition for meaning, but need not be sufficient). Neither is it meant as a repudiation of semantic externalism. Externalists believe that meaning is (completely or at least partly) determined by factors external to the speaker. But even they will have to admit that only conscious, intentional agents can count as speakers. Externalism claims that the conditions which determine the meaning of a term are not (merely) subjective – but this is wholly compatible with the fact that these terms can only exist as carriers of meaning if there are conscious beings capable of mental representation.

But while the capacity for mental representation is necessary for linguistic signs, it is not sufficient, since clearly there are beings who are capable of mental representation but not of language, e.g. most animals and very young children. What distinguishes language-users from them is their capacity for *conceptual* mental representation. Only conceptual mental contents are communicable<sup>21</sup>: if something can be said, what we

<sup>19</sup> Cf. Grice's distinction between natural and non-natural meaning in Grice (1957).

<sup>20</sup> It is a fairly common position in philosophy of language to regard mental meaning as original and linguistic meaning as derivative. For example, Aristotle (1963) says in *De Interpretatione* that 'spoken sounds are symbols of affections in the soul' (16a); Locke (1959) declares at the beginning of his treatise on meaning in the *Essay* that 'words, in their primary or immediate signification, stand for nothing but the ideas in the mind of him that uses them.' (b. III, ch. II, 2); Husserl, in his *Logical Investigations*, says that '[Expressions] serve the hearer as signs of the "thoughts" of the speaker, i.e. of his sense-giving inner experiences, as well as of the other inner experiences which are part of his communicative intention.' (Husserl 1970, 277); Peirce claims that the significate (which he calls 'interpretant') is 'a sort of idea' (2.228). Davidson, finally, remarks at the beginning of his paper on *Thought and talk*: 'The dependence of speaking on thinking is evident, for to speak is to express thoughts.' (Davidson 2001, 155)

<sup>21</sup> This also explains why it is common in the literature on ineffability to say that no concepts apply to ineffable things (Yandell 1975, 172 or Kellenberger 1979, 309), or that the ineffable is transcategorical (Hick 2000, 35) or unrepresentable (Kukla 2005, 135ff.): what is meant is that the ineffable is mentally represented, but not conceptualized.

mean is that it can be the content of a linguistic utterance, and these contents are propositions which are conceptual in nature. The ability to use linguistic signs presupposes the ability to form concepts – no language without concepts.<sup>22</sup> For example, the word ‘lemon’ is a sign that refers to no individual lemon in particular, but potentially to any lemon there is, ever was or will be. What these lemons all have in common are the particular characteristics that make them lemons – but this just means that they are all instantiations of the concept *lemon*. Calling them ‘lemons’ only makes sense if I have already grasped the concept of a lemon – otherwise the sign would not be understandable as a sign for lemons. So, if something can be said, it must be conceptual in nature, and if mental representations are conceptual, they can in principle be communicated. Vice versa, what we cannot conceptualize, we cannot say either.

It follows that strong ineffability is the inability to form a mental significate that represents the state of affairs one is trying to express, or that one is unable to conceptualize a given mental representation. If something is strongly ineffable, we are unable to give a sign the meaning required to express what we want to express. This implies at least three things: (a) strong ineffability is the wider term: if something is strongly ineffable, it is also weakly ineffable, and necessarily so; but the reverse doesn’t hold. Strong ineffability is necessary weak ineffability. If something is inexpressible in some language due to some contingent deficiencies of the language itself, it is weakly ineffable. But if it cannot be in principle expressed because no sign could ever carry the meaning that we want to express, it is strongly ineffable. While weak ineffability is language ineffability, strong ineffability is mind ineffability. (b) Weak ineffability is in principle removable by introducing new concepts to your language, strong ineffability is not (at least not for us), since it is not possible to form the relevant concepts in the first place. Any weak ineffability can in principle be removed, since all that is lacking is a term which corresponds to our concept, so that for every weakly ineffable proposition, an alternative language is conceivable in which the proposition can be expressed (even if no natural language exists that does the job). Strong ineffabilities are incurable because they do not stem from linguistic deficiencies but from mental ones; weak ineffability is relative to a language, strong ineffability is relative to a type of cognitive system. (c) The boundaries of conceptual representation of different cognitive systems are contingent. If we are unable to give a sign the required meaning, this does not mean that no sign at all could ever have this meaning. So, whether something is incommunicable or not depends on the particular type of cognitive system in question and its conceptual capacities. The bounds of language are the bounds of *our* language after all.

**Funding Information** Open Access funding provided by Projekt DEAL.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory

<sup>22</sup> Note that this does not necessarily imply the reverse: no concepts without language. – Also, my point doesn’t depend on any particular theory of concepts. All that matters is that concepts are constituents of thoughts, have a certain degree of generality, and categorize objects.

regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Alston, W. (1956). Ineffability. *Philosophical Review*, 65, 506–522.
- Aristotle (1963). *Categories and De Interpretatione*. Trans. by J.L. Ackrill, Oxford: Oxford University Press.
- Bennett-Hunter, G. (2014). *Ineffability and religious experience*. London: Pickering and Chatto.
- Carnap, R. (1937). *Logical syntax of language*. London: Routledge.
- Chomsky, N. (1957). *Syntactic structures*. The Hague: Mouton.
- Cummins, R. (1989). *Meaning and mental representation*. Cambridge: MIT Press.
- Davidson, D. (2001). *Inquiries into truth and interpretation*. Oxford: Oxford University Press.
- Grice, H. P. (1957). Meaning. *Philosophical Review*, 66, 377–388.
- Henle, P. (1949). Mysticism and semantics. *Philosophy and Phenomenological Research*, 9, 416–422.
- Hick, J. (2000). Ineffability. *Religious Studies*, 36, 35–46.
- Hofweber, T. (2005). Inexpressible properties and propositions. In D. Zimmerman (Ed.), *Oxford studies in metaphysics Vol. 2* (pp. 155–206). Oxford: Oxford University Press.
- Husserl, E. (1970). *Logical investigations*. Trans. by J.N. Findlay, London: Routledge.
- Jonas, S. (2016). *Ineffability and its metaphysics. The unspeakable in art, religion, and philosophy*. London: Palgrave.
- Kellenberger, J. (1979). The Ineffabilities of mysticism. *American Philosophical Quarterly*, 16, 307–315.
- Kukla, A. (2005). *Ineffability and philosophy*. London: Routledge.
- Lewis, D. (1970). General semantics. *Synthese*, 22, 18–67.
- Lewis, D. (1975). Language and languages. *Minnesota Studies in the Philosophy of Science*, 7, 3–35.
- Locke, J. (1959). *An essay concerning human understanding*. New York: Dover Publications.
- Lyons, J. (1981). *Language and linguistics*. Cambridge: Cambridge University Press.
- Moore, A. W. (1997). *Points of view*. Oxford: Clarendon.
- Moore, A. W. (2003). Ineffability and nonsense. *Proceedings of the Aristotelian Society*, 77, 169–193.
- Ninan, D. (2010). Semantics and the objects of assertion. *Linguistics and Philosophy*, 33, 355–380.
- Peirce, C. S. (1931). In C. Hartshorne & P. Weiss (Eds.), *Collected papers. Volume I and II: Principle of philosophy. Elements of logic*. Cambridge: Harvard University Press.
- Rabern, B. (2012). Against the identification of assertoric content with compositional value. *Synthese*, 189, 75–96.
- Shaw, J. R. (2013). Truth, paradox, and ineffable propositions. *Philosophy and Phenomenological Research*, 86, 64–104.
- Spackman, J. (2012). Expressiveness, ineffability, and nonconceptuality. *The Journal of Aesthetics and Art Criticism*, 70, 303–314.
- Speaks, J. (2014). Theories of meaning. *The Stanford Encyclopedia of Philosophy* (Fall 2014 Edition), Edward N. Zalta (ed.), URL = <<http://plato.stanford.edu/archives/fall2014/entries/meaning/>>.
- Wittgenstein, L. (1922). *Tractatus Logico-Philosophicus*. London: Routledge & Kegan Paul.
- Wittgenstein, L. (2009). *Philosophical investigations*. Oxford: Blackwell.
- Yandell, K. (1975). Some varieties of ineffability. *International Journal of Philosophy of Religion*, 6, 167–179.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.