

The Building Blocks of Scientific Knowledge: Neurath and Carnap on Protocol Sentences

Joshua Anderson

Department of History and Philosophy

Virginia State University

1 Hayden Drive, Box 9070, Petersburg, VA 23806, USA

Email: janderson@vsu.edu

ORCID: 0000-0002-3621-9771

Abstract:

In this paper, I discuss the protocol sentence debate between Otto Neurath and Rudolf Carnap. Despite their differences, both Carnap and Neurath maintain that protocol sentences play an essential role in the goals and purpose of philosophy. Particularly, they believe that philosophy should be focused on trying to advance scientific knowledge, and both thinkers see protocol sentences as the building blocks of science. The paper will consist of three main parts. First, I will present Otto Neurath's account of protocol sentences as explained in his essay, "Protocol Sentences." The second portion of the paper will explain Rudolf Carnap's account of protocol sentences as presented in his essay, "On Protocol Sentences." The third portion of the paper will compare and contrast the two versions of protocol sentences.

Keywords: Rudolf Carnap, Otto Neurath, Protocol Sentences

1. Introduction

One of the main goals of Logical Positivism and the Vienna Circle was to alter the trajectory of philosophy. They sought to move philosophy away from what they believed were unverifiable metaphysical speculations and make it more "scientifically respectable." Indeed, they sought to ground science in a fully verifiable empiricism. An important discussion arose among them regarding protocol sentences. Protocol sentences articulated and described the immediate experiences that could serve as the ultimate empirical ground of knowledge, in general, and scientific knowledge, in particular. Otto Neurath and Rudolf Carnap had distinct accounts of protocol sentences and a vigorous debate ensued.

It seems that for both Carnap and Neurath protocol sentences play an essential role in their overall project of how philosophy should proceed. Particularly, both thinkers are trying to advance scientific knowledge, and to this end, both thinkers see protocol sentences as the building blocks of science. Protocol sentences provide the justification for choosing particular sentences to be included within the domain of science. Protocol sentences also provide the content of the sentences within the language of science—this second point may not be necessarily true for Neurath but is certainly true of Carnap.

In this paper, I discuss this debate by comparing and contrasting the main features of Neurath's and Carnap's understanding of protocol sentences. The paper will proceed as follows: First, I present Otto Neurath's account of protocol sentences as explained in his essay, "Protocol Sentences" (1959). Next, I explain Rudolf Carnap's account of protocol sentences as presented in his essay, "On Protocol Sentences" (1987). Finally, I compare and contrast the two versions of protocol sentences, concluding that Carnap provides a preferable version of protocol sentences.

2. Neurath and Protocol Sentences

A critical point to understand is that, for Neurath, one already exists within a language. The language one already speaks must be acknowledged and addressed. One cannot step outside the language in order to begin again with a more perfect or ideal language. Neurath gives the following analogy to explain his point: "We are like sailors who must rebuild their ship on the open sea, never able to dismantle it in dry-dock and to reconstruct it there out of the best materials" (Neurath 1959, 201). All that can be done is to, first of all, rid the language of as many metaphysical aspects as possible and then to proceed scientifically to make the language as clear as possible. More importantly, for Neurath, the ultimate goal is to create a unified language of all of science—from physics to sociology and biology to economics.

Within the unified language of science, protocol sentences are a special case of factual sentences. For Neurath, protocol sentences are very precise nested sentences that include repeated use of a personal noun. "A complete protocol sentence might, for instance read: 'Otto's protocol at 3:17 o'clock: [At 3:16 o'clock Otto said to himself: (at 3:15 o'clock there was a table in the room perceived by Otto)]'" (Neurath 1959, 202).

Two points need to be made when analyzing the example of Neurath's version of protocol sentences because they become essential aspects of any protocol sentence for Neurath. The first point is that it involves spatio-temporal phenomena. The second point is that protocol sentences are inter-subjective, at least in principle. Point one is important for Neurath's view of science. Science talks about the actual physical world, thus when one experiments scientifically the experiments are judged in relation to physical spatio-temporal phenomena, and one talks of spatio-temporal phenomena using the protocol sentences. Thus, the above protocol sentence is implying there is a table.

The second point is more interesting. The repeated use of a personal noun is extremely important. For Neurath, every language is inter-subjective. Thus, "[t]he protocols of one moment must be subject to incorporation in the protocols of the next, just as the protocols of A must be subject to incorporation in the protocols of B" (Neurath 1959, 205). The use of a personal noun emphasizes the fact that one stands in the same relation to any protocol sentence. The upshot is that "[o]ne can distinguish an *Otto-protocol* from a *Karl-protocol*, but not a protocol of one's own from a protocol of others. The whole puzzle of *other minds* is thus resolved" (Neurath 1959, 206).

Before summing up Neurath's position a few issues need to be clarified. First, Neurath believes that one already exists within a language that is different from the universal language of the sciences.

After purging this language of as many metaphysical aspects as possible one is left with a “universal slang.” The slang can never really be done away with and aspects of it will always exist. The only difference between the universal slang and the universal language of science is perhaps precision, but science will always rely in some part on the slang—this goes back to Neurath’s analogy of rebuilding the boat.

The second issue, that perhaps has not been made clear, is that for Neurath no statements, or sentences, are incorrigible. Like many within the Vienna Circle, Neurath is deeply concerned with verification. For Neurath, part of a statement’s verifiability is the ability of one to reject the statement from the language.

Ultimately, Neurath looks to how science is done to guide his philosophy and that includes his understanding of protocol sentences. There are some key features of Neurath’s view. One feature is that there is no private language and that language is always inter-subjective. A second feature is that language is necessarily vague and vagueness will always be part of any language. A third distinctive feature is fallibilism; that is to say that all statements of the language are revisable or rejectable. Finally, it should be pointed out that while any statement can be discarded, there is nothing that could force one to discard any particular statement. Neurath looks to scientific and social values to guide one’s decisions in accepting or rejecting particular statements. It is this dependence on scientific and social values that help structure Neurath’s conception of a protocol sentence.

3. Carnap and Protocol Sentences

Carnap’s conception of protocol sentences is quite different from Neurath’s. Carnap refers to his protocol sentences as outside the system, while those of Neurath are internal to it. It appears that, for Carnap, protocol sentences are responses—physical or verbal—to the world. Carnap gives the example of a machine. The machine gives certain signals based on the weather. The signals are then translated into the physical language based on particular rules that Carnap calls T-rules or rules of translation. The same situation occurs with a person. The person gives vocal or written “signals” that are then translated into the physical language via the T-rules. To be clear, it is the “signals” that are the protocol sentences for Carnap.

Unlike Neurath, protocol sentences are incorrigible for Carnap. The incorrigibility of the protocol sentence makes sense because they are merely reactions to an environment. It is not until the protocol sentences have been translated via the T-rules that they have an understandable meaning. Further, just because one cannot revise a particular protocol sentence does not mean that the particular protocol sentence will be translated as true. Finally, T-rules translate protocol sentences into sentences about the environment in general, which Carnap calls T-sentences. However, Carnap points out that “[o]bservation teaches that with the use of T-rules the acquired T-sentences are not very reliable” (Carnap 1987, 460). It is for this reason that Carnap introduces a second type of translation rule.

The second type of translation rule Carnap calls B-rules, and they translate protocol sentences into B-sentences. B-sentences do not speak about the environment the way T-sentences do. Rather, they speak about “the momentary state of the body B of the person in question. Thus, instead of translating a protocol, via T-rules to the corresponding T-sentence: “It is raining”, one would translate the protocol, via B-rules to the corresponding B-sentence: “The body β is rain-observing”.

From the B-sentence we can infer with probability but not with certainty the corresponding T-sentence...; the probability increases if it is known to us that the environment of β is in a “normal” state. The inference from B-sentence to the T-sentence is nothing but the usual inference from the effect to a probable cause (Carnap 1987, 461).

However, it is to be remembered that protocols are taken as they are presented without making a selection. It is not until the protocol is translated into the system language via B-rules and T-rules that one can make sense of the rules.

If one comes across conflict or inconsistency one must adjust the B- or T-rules. Again, translating a sentence does not mean that the particular sentence is true, just like translating between two natural languages. A false statement in German can be translated to a correspondingly false statement in English without doing violence to either statement or forcing one to hold inconsistent beliefs. Also, “[s]ince [protocol sentences] are to be understood at the outset as processes and not as sentences in the language, then there is literally no contradiction. There can be a contradiction only between [...] sentences into which we translate the [protocol sentences]” (Carnap 1987, 461).

Still, the status of the B- and T-rules remains a bit ambiguous. It is not clear whether the rules of translation are actually part of the language system. If the translation rules are part of the system, then to change the translation rules is to change the system language. If the translation rules are not part of the language system, then it is an open question as to their status and how they relate to the system language.

After explaining his view of protocol sentences Carnap concludes by explaining how exactly the protocol sentences fit into his overall view of how science works. First, there exists a system language—for Carnap this is the language of science. Second, there are universal sentences—for example, laws of nature—and concrete sentences. “[o]utside the system language there are signals which are understood as ‘protocol sentences’ of a ‘protocol language’” (Carnap 1987, 462). B- and T-rules are set up to translate the protocol sentences into concrete sentences of the system language. Then using the concrete sentences more concrete sentences and universal system sentences can be constructed or derived.

4. Comparing and Contrasting

While I do not agree entirely with Carnap’s view, I do believe that his position is more plausible than Neurath’s. Though Neurath’s goal is to work with the language one already has, his view of protocol sentences seems to be overly burdensome and unnatural. Neurath does not seem

to claim that people should go around speaking in protocol sentences. Yet, if one does not speak, at least at certain times, in protocol sentences it is uncertain what the point is. I suppose that the point for Neurath is to be clear. When one states there is a table in the room what is really meant is something similar to the example of the protocol sentence given above—viz. ‘Otto’s protocol at 3:17 o’clock: [At 3:16 o’clock Otto said to himself: (at 3:15 o’clock there was a table in the room perceived by Otto)]’” (Neurath 1959, 202).

I also agree with Carnap that Neurath’s version of a protocol sentence “with its three nested components is [not] practical; it has the defect, from the point of view of syntax, that a sentence which refers to another contains the other as a clause” (Carnap 1987, 465). I do think that Neurath’s move of introducing the personal noun as an integral part of the protocol sentence is quite brilliant. By eliminating references to ‘I’ or ‘mine’, the problem of other minds is eliminated, at least in a certain respect. Yet, it does not seem one stands in the same relation to all protocols. Neurath seems to assert that it is a mere fiction that one prioritizes one’s own statements. If Neurath is correct then the problem of other minds is solved, though I believe he needs to be more explicit as to why it is that one’s own protocols are not different from any other protocol.

I take it that Neurath’s only defense is that reference to an ‘I’ or ‘others’ imports problematic metaphysical notions. Therefore, when one removes the problematic metaphysical aspects from the universal slang one can only refer to personal nouns. If one wished to be more specific a space-time co-ordinate or picture could be used as a reference as opposed to a name.

One final point needs to be made before moving on to Carnap. Neurath is able to keep a physical world, though not in a robust metaphysical sense, because language is a physical act. Also, because language cannot be private and can only be inter-subjective, again not only is the problem of other minds solved, but the world is reaffirmed.

I find Carnap’s position more plausible. Though on the face of it appears to be rather complex, on reflection it does seem to be how science is actually done. For example, when one looks at a thermometer all that is being signaled by the thermometer is the mercury has risen to a line with a number written next to it. It is only through translating the signal of the thermometer that sense can be made of the thermometer’s signal.

If it is true of thermometers, then it naturally there may be a correlation with how people speak. Of course, in actual practice the translation from the protocol language to the system language may not be done in separate steps as Carnap has explained it. There still remains the problem of how the translation rules relate to the system language. It is not clear if the translation rules are part of the language. If they are not part of the language Carnap does not explain what their status is. If the translation rules are part of the language, then to change a translation rule changes the system language and would thus make it a different language than the language one was speaking originally.

One could reject Carnap’s account of protocol sentences because the protocol sentences are part of what Carnap calls a protocol language, which seems essentially private. It is generally taken that an essentially private language, is at best problematic, but more likely impossible (See

Wittgenstein 2009 and Kripke 1982). Also, what distinguishes “language” acts from any other acts? On Carnap’s account these are only pseudo-questions since one needs to adopt a specific language to adjudicate questions of this sort.

Another aspect of Carnap’s account that I find appealing is that questions of meaning or connection with the external world do not or cannot occur. When one accepts a language, the logical syntax fixes the meaning of terms and sentences. Thus, meaning is not dependent on relation to an external world. To speak of the truth of a particular statement in relation to its connection to something external is a meaningless pseudo-question.

On the other hand, if meaning and connection to the external world come apart, it appears that protocol sentences would not even be necessary for Carnap’s account. Carnap could accept that. Carnap has chosen a system language that is dependent on protocol sentences to provide the content of certain concrete sentences within the language. Moreover, Carnap realizes that others may not wish to do so. Carnap refers to his position as a proposal. He is not affirming how things *are*. Carnap may not even be claiming how things *should be*.

Carnap has chosen his particular view for pragmatic reasons. He believes that his view can help explain how science is done, and how science can continue to progress. Carnap would not even go so far as to say that his account is the best, or simplest, et cetera. For Carnap, ultimately, there is no fact of the matter regarding which language one should choose. Indeed, it is actually this lack of dogmatism that I find to be one of the most appealing features of Carnap’s view.

While Neurath’s account appears to be simpler because it deals with only one language, this aspect actually makes the view more burdensome. Though Carnap’s view has possibly three distinct levels or aspects—protocols, translation rules, and the system language—the explicit compartmentalizing makes his view more understandable. It seems that what is explicit in Carnap is implicit in Neurath. Carnap has the “advantage that it places no demands regarding protocolling” (Carnap 1987, 469). Conversely, Neurath’s understanding of protocol sentences is highly demanding.

Finally, Carnap admits his arbitrariness. He does not claim to have all the answers. While Neurath is searching for *the* universal language of the sciences, Carnap is satisfied with much less. Carnap wants to become clear on what is being said. Carnap merely seeks to point out the assumptions that are being made, not to condemn the assumptions necessarily, just to make those assumptions explicit.

5. Conclusion

I have presented above what I believe are the essential features of Neurath’s and Carnap’s account of protocol sentences. During the discussion I have also addressed issues within Neurath’s and Carnap’s philosophy more generally. While I have tried to limit those forays, I do believe they were necessary. Ultimately, I have found Carnap’s view more plausible for two reasons. First, I find Carnap’s project generally, not just his position on protocol sentences, more appealing than Neurath’s. A complete discussion of why I find Carnap’s view generally more plausible extends

beyond the scope of this paper, though I have at times hinted at it. The second reason I find Carnap's account of protocol sentences more plausible is because he puts virtually no constraints on the protocol sentence itself.

It seems to me that Neurath's presuppositions force him to construct overly burdensome and complex protocol sentences that are completely artificial. Carnap's version comes across as more natural and intuitive. Carnap's account does become more complex with the introduction of T-rules, B-rules, T-sentences, B-sentences, and system language. Yet, the base unit, the protocol sentence is left clear. It is then up to one's assumptions and linguistic framework to make sense of the protocol sentence. Neurath, on the other hand, forces the assumptions and values in at the very beginning by constraining the protocol sentences themselves.

References:

- Carnap, Rudolf. "On protocol sentences." *Noûs* 21. 4 (1987): 457-470.
- Kripke, Saul. *Wittgenstein on rules and private language: An elementary exposition*. Cambridge, Mass.: Harvard University Press, 1982.
- Neurath, Otto. "Protocol sentences." *Logical positivism*. Ed. A.J. Ayer. New York: The Free Press, 1959. 199-208.
- Wittgenstein, Ludwig. *Philosophical investigations*. Revised 4th ed. Malden, MA: Blackwell Publishing, 2009.